

25133
nature

Vol 290 No 5801 5 March 1981 £1.25 \$3.00

PROPAGATION OF EXPLOSIONS

153

153

(Vol 290)

nos. 5801-5809

(March - April)

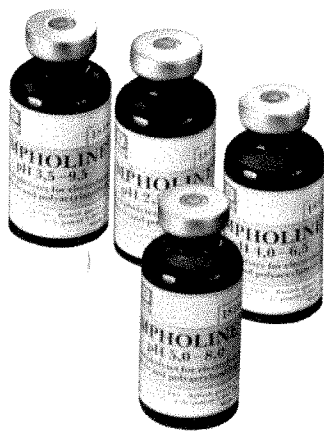
(4-806 p.)

(1981)

LKB ANNOUNCES MORE.

MORE CONVENIENCE

New pre-blended Ampholine[®] carrier ampholytes for electrofocusing are available in three narrow and one wide pH range. Now there's no need to spend time mixing to achieve these specific ranges. Just select the pre-blended Ampholine you need (pH 2.5 to 4.5, 4.0 to 6.5, 5.0 to 8.0 and 3.5 to 9.5). They're designed to give the same linear, virtually drift-free pH gradient achieved using standard Ampholine with LKB Agarose-EF or acrylamide gels.

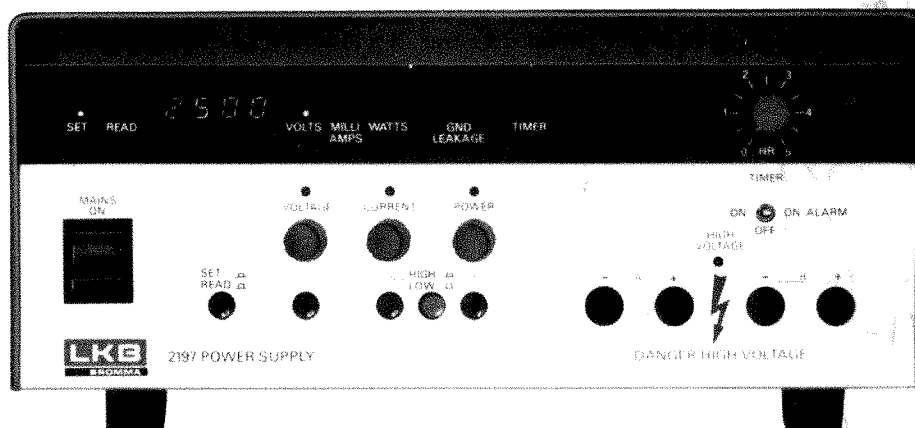
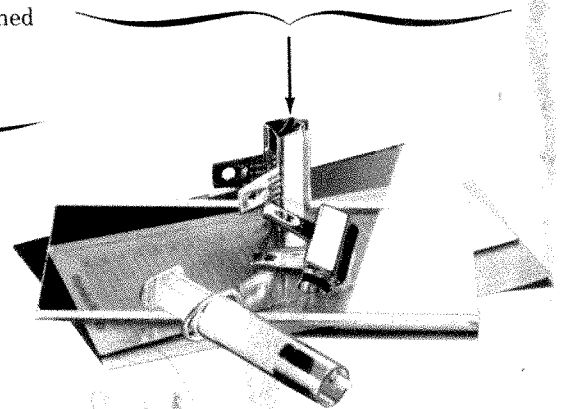


MORE POWER

The new Model 2197 Power Supply for electrofocusing and electrophoresis offers top-of-the-line features at a very affordable price. The Model 2197 delivers up to 2500V for high-resolution electrofocusing and its 250mA output makes it useful for other electrophoretic techniques. Choose constant power/constant voltage/constant current/time control modes to suit your application. Automatic crossover between running modes compensates for varying gel conditions. Built-in timer with alarm and automatic shutoff allows for time control. With accurate control of four parameters, you can achieve maximum resolution and reproducibility in all your electrofocusing and electrophoresis experiments. Bright LED display lets you monitor performance easily and the Model 2197 is designed to comply with all relevant safety standards.

MORE ECONOMY

LKB's new capillary gel casting system easily produces gels only one quarter the thickness of standard ones. You save substantially on reagent use and you'll be able to employ higher voltage for decreased run times because the new thinner gels dissipate heat more efficiently. The capillary gel casting kit works with agarose and acrylamide gels. Combine it with the new pre-blended Ampholine and the 2500 volt power supply to significantly reduce your run times while cutting costs. Ask your local LKB representative for more information on these new products and the complete line of LKB instruments and consumables. Depend on the leader in electrofocusing to give you more.



LKB

Head office:
LKB Produkter AB
Box 305
S-161 26 Bromma, Sweden
Tel. 08/98 00 40, telex 10492

UK sales office:
LKB Instruments Ltd.
232 Addington Road
South Croydon, Surrey CR28YD
England

Other sales offices in:
Athens (for the Middle East),
Copenhagen, Ghent,
The Hague, Hong Kong,
London, Munich, Paris,
Rome, Turku, Vienna

Circle No. 20 on
Reader Enquiry Card.

nature

Volume 290 No.5801 5 March 1981

NEWS

Pope John Paul II in Hiroshima	1	Soviet research: More home growth	5
Eruptions in space	1	German nuclear power: Closures and delays	6
Questions answered	2	Human growth hormone: Pituitary slump	6
Open row about joint space project: European space agency plans protest	3	University of London: Storm ahead	7
Gloom at prospect of cancellation	3	Soviet activists honoured	7
European fusion: JET's ambition	4	Chemical and nuclear weapons: Scientists speak out	7
JET upgraded	4	Correspondence: Data bank (M.O. Dayhoff, R.M. Schwartz, H.R. Chen, L.T. Hunt, W.C. Barker and B.C. Orcutt); Antipodean 2,4,5-T (P. Hall, B. Selinger, B. Field and C. Kerr); Cell contamination (G.E. Moore and L.K. Woods)	8
United States budget: More cuts	4		
UK research councils: Allen accused	5		

NEWS AND VIEWS

Ice and ocean: Starley L. Thompson and Stephen H. Schneider on carbon dioxide and climate	9	γ physics — from bottom to top: Roger Cashmore surveys the field	13
Active galaxies: Martin Elvis and Andrew S. Wilson describe new radio and X-ray observations	10	Molecular hydrogen in T Tauri: Stephen P. Maran comments on observation from the International Ultraviolet Explorer	14
Designer genes: a correspondent on the problems of expression of ectopic genes in mammalian cells	11	Planarian neoblasts: J. Baguña and J.M.W. Slack discuss regeneration in planaria	14
Astrophysics and particle physics come together: Frank E. Close reports a recent conference	12	A carbohydrate signal for intracellular transit: Mike Geisow explains how the signal function was verified	15
		Beyond the giga seal: Robert McBurney on new techniques for monitoring membrane ion channels	16

ARTICLES

C.C. Covey and G. Schubert	Mesoscale convection in the clouds of Venus	17	S.A. Liebhaber, M. Goossens and Y.W. Kan	Homology and concerted evolution at the $\alpha 1$ and $\alpha 2$ loci of human α-globin	26
D.V. Goeddel, D.W. Leung, T.J. Dull, M. Gross, R.M. Lawn, R. McCandliss, P.H. Seeburg, A. Ullrich, E. Yelverton and P.W. Gray	The structure of eight distinct cloned human leukocyte interferon cDNAs	20	S.C. West, E. Cassuto and P. Howard-Flanders	Homologous pairing can occur before DNA strand separation in general genetic recombination	29

LETTERS

A. Brown, C. Jordan, T.J. Millar, P. Gondhalekar and R. Wilson	H₂ emission in the EUV spectrum of T Tauri and Burnham's nebula	34	S.R. Moore and F.J. Weinberg	High propagation rates of explosions in large volumes of gaseous mixtures	39
E. Scalise Jr and M.A. Braz	A water vapour maser in the Large Magellanic Cloud	36	G. Davies	Decomposing the IR absorption spectra of diamonds	40
B. Singh, R. Prakash and H. Singh	Discrete chorus emissions observed at a low latitude ground station	37	D.E. Fisher	Quantitative retention of magmatic argon in a glassy basalt	42

Cover: Infinitesimal shear, double exposure interferogram showing temperature near a radiation-ignited model house (Weinberg & Wong, Imperial College). S.R. Moore and F.J. Weinberg speculate on the role of radiation-ignition in large-scale explosions on page 39.

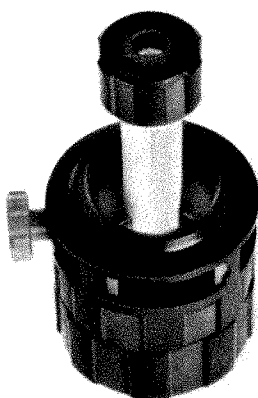
M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010. © 1981 Macmillan Journals Ltd.

Contents continued overleaf

The Complete Column

- Everything you need comes in a convenient do-it-yourself kit
- Applicable for any bed height from 5 to 95 cm
- The same low cost for both $\varnothing$ 16 and 26 mm I.D.



- Assembles in less than 5 minutes without any extra tools
- Packing cylinder and reservoir available as accessory
- Sample applicator included



LKB

Head office:
LKB-Produkter AB
Box 305
S-161 26 Bromma, Sweden
Tel. 08-98 00 40.

Main US sales office:
LKB Instruments, Inc.
12221 Parklawn Drive
Rockville, MD 20852
Tel. 301-881 25 10.

UK sales office:
LKB Instruments Ltd
232 Addington Road
S. Croydon, Surrey CR2 8YD
England
Tel. 01-852 9999. Telex 954411

Other sales offices in:
Athens (for Middle East),
Copenhagen, Ghent,
The Hague, Hong Kong,
Munich, Paris, Rome,
Tokyo, Vienna.

LETTERS

E.G. Garrison, R.M. Rowlett, D.L. Cowan and L.V. Holroyd	ESR dating of ancient flints	44	T. Watanabe and J. Ohara	Functional nuclei of LPS- nonresponder C3H/HeJ mice after transfer into LPS-responder C3H/HeN cells by cell fusion	58
M.W. Pitts	Stones, pits and Stonehenge	46	C. Martinez-A and A. Coutinho	B-cell activation by helper cells is a two-step process	60
K. Brander	Disappearance of common skate <i>Raja batis</i> from Irish Sea	48	A.T.J. Bianchi, H. Hooijkaas, R. Benner, R. Tees, A.A. Nordin and M.H. Schreier	Clones of helper T cells mediate antigen-specific, H-2-restricted DTH	62
M.J.A. Simpson, A.E. Simpson, J. Hooley and M. Zunz	Infant-related influences on birth intervals in rhesus monkeys	49	M.G. Rosenfeld, S.G. Amara, B.A. Roos, E.S. Ong and R.M. Evans	Altered expression of the calcitonin gene associated with RNA polymorphism	63
J. Skelly, C.R. Howard and A.J. Zuckerman	Hepatitis B polypeptide vaccine preparation in micelle form	51	A.L.M. Bothwell, M. Paskind, R.C. Schwartz, G.E. Sonenshein, M.L. Gefter and D. Baltimore	Dual expression of λ genes in the MOPC-315 plasmacytoma	65
P.J. Cowen, A.R. Green, D.J. Nutt and I.L. Martin	Ethyl β-carboline carboxylate lowers seizure threshold and antagonizes flurazepam-induced sedation in rats	54	L.M. Silver and K. Artzt	Recombination suppression of mouse <i>t</i>-haplotypes due to chromatin mismatching	68
F. Scaravilli and J.M. Jacobs	Peripheral nerve grafts in hereditary leukodystrophic mutant mice (twitcher)	56			

BOOK REVIEWS

Alastair Hay	The Superpoison (by T. Margerison, M. Wallace and D. Hallenstein)	71	M.J. Morgan	Essay on Mind (by D.O. Hebb)	73
A.D.R. Phelps	Plasma Physics for Nuclear Fusion (by Kenro Miyamoto). Fundamentals of Plasma Physics (by V.E. Golant <i>et al.</i>)	72	Adrian Friday	Comparative Biology and Evolutionary Relationships of Tree Shrews (edited by W.P. Luckett)	74
Alan E.H. Emery	The Principles of Human Biochemical Genetics, 3rd Edn (by Harry Harris)	72	Paul Davies	Quantum Field Theory (by Claude Itzykson and Jean-Bernard Zuber)	74

MISCELLANY

New on the market	xix	Classified advertising	xxii
Announcements	xxi		

GUIDE TO AUTHORS

● Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.

● Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.

● Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.

● 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov. (1976).

LIGHT MICROSCOPY '81 — THE STATE OF THE ART

International Conference and Exhibition on Light Microscopy

Imperial College, London SW7 7-9 July, 1981

The Exhibition, is organised by the Royal Microscopical Society in association with 'Nature'.

The Conference not only marks the inauguration of the Light Microscopy Section of the Royal Microscopical Society, but establishes a landmark in the field by providing a definitive statement of the more important aspects of light microscopy as it is now and of significant developments.

The first day's programme is devoted to the instrument itself. The second day is concerned with imaging techniques, and "light" microscopy is interpreted liberally. The third day will deal with two aspects of measurements on images — photometric (in transmission, emission and reflection) and geometric.

A trade exhibition will be held in conjunction with the meeting. In it, manufacturers will show equipment at the forefront of development and pertinent to the meeting programme. The exhibition will be open, *free of charge*, to both meeting registrants and non-registrants. Times when non-registrants may visit the exhibition are given below. (A registration form is inserted in this Journal.)

PROVISIONAL PROGRAMME

Tuesday, 7 July — The Instrument

Survey Paper — Modern Microscope Design

Modern Microscope Design — single or multi-purpose routine or research work

The Design of Microscopes for Fluorescence Techniques

Photomicrographic Systems

The Design of Stereo Microscopes

Trends in Microscope Objective Design

A New Optical Design for Light Microscopes

Role of Specialist Microscope Manufacturer in the 1980's

Wednesday, 8 July — Developments in Imaging Techniques

Review — Imaging Methods

Holographic Methods

Acoustic Microscopy

Laser-Scanning Microscopy

A Combined Light Microscope-Scanning Electron Microscope

Possibilities and Limitations of Interference Microscopic

Phase Analysis

Developments in Polarised Light Microscopy

Professor H Hasselmann, Inst. für Wissenschaftliche
Mikroskopie der Universität, Tübingen
Dr Klaus Weber, Carl Zeiss (Oberkochen)

Dr A Thaer, Ernst Leitz Wetzlar GmbH
Dr K Schindl, C Reichert Optische Werke AG
Dr J T Watts, Wild-Heerbrugg
Dr Z Wakimoto, Nippon Kogaku K.K., Tokyo
Mr A Taira, Olympus Optical Co. Ltd, Tokyo
Mr J H Bassett, James Swift & Sons Ltd

Professor G Nomarski, CNRS, France
Dr R W Smith, Imperial College, London
Dr Wickramasinghe, University College, London
Dr C Shepherd, University of Oxford
Professor J S Ploem, University of Leiden
Dr J Bergner, Carl Zeiss Jena

Mr D A Hemsley* and Mr P C Robinson**
*Institute of Polymer Technology, Loughborough
**North Staffordshire Polytechnic

Thursday, 9 July — Measurements

Review — Quantitative Automated Microscopy

Photometric Analysis of Reflectance and Fluorescence
emission spectra

Microphotometry Instrumentation

Quantitative Interference Microscopy

Fluorescence-activated Cell Sorting

Systems for Acquisition of Microscopical Information for
Image Analysis

Image Analysis in Pathology

Mathematical Morphology Applied to Image Analysis of

Cytological Specimens

The Use of a TV-Image Analysis Computer System (LEYTAS)
in Clinical Cytology

Fully Automated Systems

Professor J S Ploem, University of Leiden
Dr H Piller, Carl Zeiss (Oberkochen)

Dr H Wasmund, Ernst Leitz Wetzlar GmbH
Dr D J Goldstein, University of Sheffield
Dr D Capellaro, Imperial Cancer Research Fund, London
Dr D Rutovitz, MRC Edinburgh

Dr J P A Baak, Delft
Dr F Meyer, Centre de Mathématique Morphologique,
Fontainebleau
Mrs A M J van Driel-Kulker, Centre de Mathématique
Morphologique, Fontainebleau
Mr A Terrell, Cambridge Instruments Ltd

EXHIBITION

In conjunction with the conference, there will be an exhibition of the latest products of the Light Microscope Manufacturers. This exhibition will be open, *free of charge*, to registrants and non-registrants. Times for non-registrants to visit the exhibition will be:

Tuesday, 7 July	—	12 noon to 1800 hours
Wednesday, 8 July	—	1100 hours to 1800 hours
Thursday, 9 July	—	1100 hours to 1500 hours

REGISTRATION AND ACCOMMODATION

Details available from the Administrator:

THE ROYAL MICROSCOPICAL SOCIETY
37/38 St. Clements, Oxford OX4 1AJ
Telephone: Oxford (0865) 48768/721081

nature

Volume 290 No.5802 12 March 1981

How true is the theory of evolution?	75	Chemical warfare: US prepares?	79
Mixed welcome for genes on Wall Street: Genentech's interferon from yeast plasmid	77	Allen's successor	79
Benefits and snags of yeast plasmids	77	Chemical industry: ICI off course	80
Thatcher applauds genes	78	Indian oil exploration: Partners sought	80
Darwin in California: No verdict	78	UK pharmaceuticals: Bad risks	81
European Space Agency: Sixes and sevens	78	Telecommunications: Bell's bother	81
One step to LEP	79	Correspondence: New IQ test? (C.R. Brand); Darwin's survival (H.W. Ball <i>et al.</i>); Conservation sites (C. Muir); Origin of cancer (D.I. Edwards; M. Woodruff)	82, 173

NEWS AND VIEWS

The intergalactic medium: Joseph Silk on the search for evidence of its existence	83	Hydrothermal activity at mid-ocean ridge axes: J.M. Edmond describes the results of deep-sea searches	87
Directing toxins to cancer cells: Sjur Olsnes assesses progress in this field	84	Monitoring Mount St Helens: M. McCormick reports on two recent meetings	88
The world's worst weeds: Robert M. May describes a programme to prevent accidental introduction	85	Will the real Grenville Orogeny please stand up?: Keith Bell on the Proterozoic in Canada	89
Geomagnetic field modelling using satellite data: David R. Barraclough reassesses the present models	86		

REVIEW ARTICLE

B. Julesz	Textons, the elements of texture perception, and their interactions	91	R.D. Evans, J.H. Harley, W. Jacobi, A.S. McLean, W.A. Mills and C.G. Stewart	Estimate of risk from environmental exposure to radon-222 and its decay products	98
-----------	--	----	--	---	----

ARTICLES

R.M. Hjellming and K.J. Johnston	Structure, strength, and polarization changes in radio source SS433	100	J.R. Nevins and M.C. Wilson	Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing	113
W.A. Hendrickson and M.M. Teeter	Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur	107			

LETTERS

M.J. Coe and D.T. Wickramasinghe	UV observations of X-ray sources 2A0311-227 and 2A0526-328	119	F.A. Shillington	Low frequency 0.045-Hz swell from the Southern Ocean	123
L. Blitz, J.W.-K. Mark and R.P. Sinha	Why is the central gas disk of the Galaxy tilted?	120	R.S. Harmon and H.P. Schwarcz	Changes of ^2H and ^{18}O enrichment of meteoric water and Pleistocene glaciation	125
G. Bonino, P.P. Lombardini, A. Longhetto and P. Trivero	Radio acoustic measurement of fog-capping thermal inversions	121	A.G. Baer	Two orogenies in the Grenville Belt?	129

Cover: The three textures on the cover appear very distinct even though they have identical third-order statistics and thus identical second-order statistics and Fourier power spectra. Experiments reviewed on pages 91 to 97 show that 'instantaneous' texture discrimination cannot be explained solely in terms of global properties of the texture and that discrimination is the result of local, conspicuous features.

M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd.

Contents continued overleaf

CURRENT CONTENTS: INDISPENSABLE!

CURRENT CONTENTS...

the best way to keep up with new
developments in your field—while
they're still new!

"By way of comment on Current Contents...your publication is indispensable around here. It comes back from circulation battered and torn -- obviously well used -- and people request references from journals we could never hope to subscribe to. Through Current Contents our library's collection is virtually unlimited.

Regarding OATS, since Current Contents generates so many reference requests, I find that OATS is really a time saver. I don't have to spend a lot of time and effort trying to locate a source for some of the more exotic titles you include in CC."

Mary Jane Campbell

Mary Jane Campbell, Librarian
Ferro Chemical
7040 Krick Road
Bedford, Ohio 44146

Current Contents, the world's most widely-used literature alerting service, displays tables of contents of the most recent issues of over 1,000 important journals and more than 700 multi-authored books. Every week top journals from all over the world are covered. . .giving you precise, but easy access to the articles you need to keep abreast of the latest developments in your field.

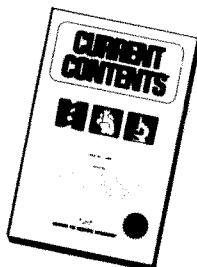
Current Contents also gives you several alternative ways to obtain any article quickly and easily if your library doesn't carry the journal. You can use *CC*'s Author Address Directory to write to the author for a reprint. Or if the article is needed immediately, you can request it from ISI's OATS® (Original Article Tear Sheet) service. For a minimal fee, OATS will have

a copy of the article in the mail to you within 24 hours after receipt of your request. You can then use the record of article requests which originated from *CC* to determine which journals you might want to subscribe to. . .or might ask your library to subscribe to. Over 300,000 researchers, scientists, and teachers rely on *CC*, coupled with OATS, to provide the journal coverage they need.

But don't just take our word for it. Why not send for a free copy of *CC* and see how it will enable you to keep up with new developments in your field — while they're still new. You have seven broad disciplinary areas from which to choose. To order your *free* issue — simply use the coupon below.

Please send me a free copy of the *Current Contents*® edition(s) I've checked below. I understand there is no cost or obligation.

- ☐ *CC*®/Life Sciences
- ☐ *CC*/Physical, Chemical & Earth Sciences
- ☐ *CC*/Agriculture, Biology & Environmental Sciences
- ☐ *CC*/Engineering, Technology & Applied Sciences
- ☐ *CC*/Clinical Practice
- ☐ *CC*/Arts & Humanities
- ☐ *CC*/Social & Behavioral Sciences



ISI® Institute for
Scientific Information®

3501 Market Street, University City Science Center
Philadelphia, Pa 19104 U.S.A.
Tel: (215) 386-0100. Cable: SCINFO Telex: 84-5305

NAME

ORGANIZATION/DEPT.

ADDRESS

CITY

STATE/PROVINCE

COUNTRY

ZIP/POSTAL CODE

PHONE

21-2510

©1981 ISI

MICRO-gTM VIBRATION ISOLATORS

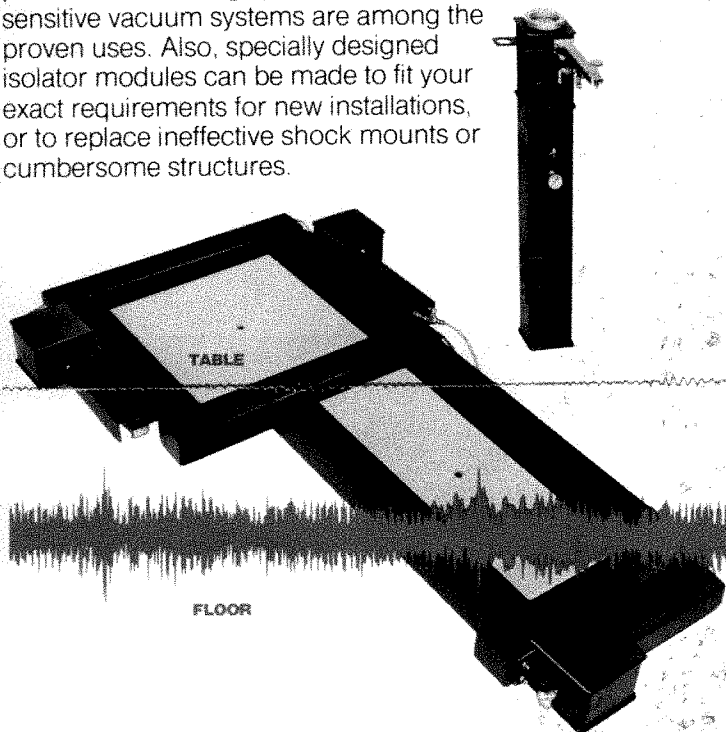
- New standards of performance
- New, exclusive isolators
- Sensible prices

IMPROVED SMALL TABLES

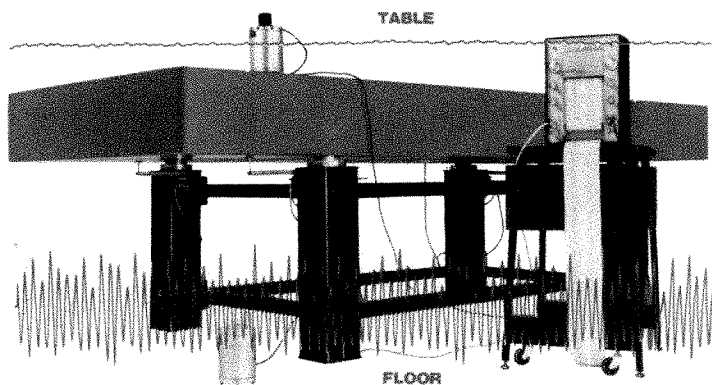
MICRO-g air isolation tables have always given exceptional performance at a good price. Now, a number of engineering improvements have been made "standard" and have increased vibration isolation significantly... about 40% better in the low frequency range where it's hard to improve any small isolator. Or, you can specify our new, optional Gimbal Piston (see below) for the best isolation in any small table system. Other improvements include stiffer, deader laminated stainless steel tops and a growing diversity of modular isolator frames that assure an available solution for your special problems.

BASE-MOUNT PLATFORMS AND SPECIAL MODULES

Our exclusive, integrated, damped platform-and-isolator design solves the problem of isolating ungainly instrument consoles and large test systems with high centers of gravity. Micro-circuitry mask aligners, electron microscopes, precision measuring systems, and extra-sensitive vacuum systems are among the proven uses. Also, specially designed isolator modules can be made to fit your exact requirements for new installations, or to replace ineffective shock mounts or cumbersome structures.

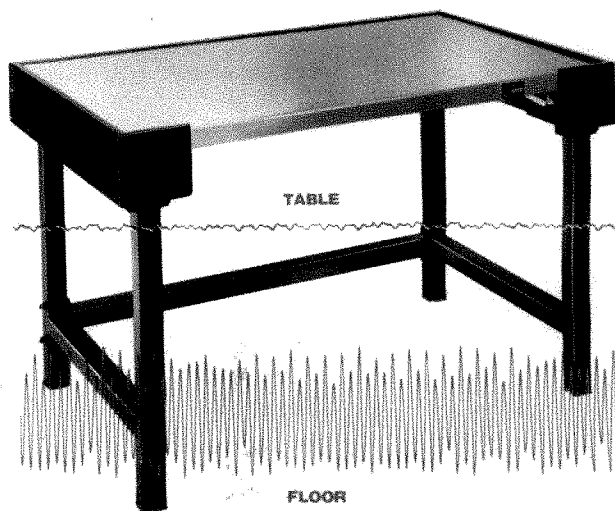


TM-5



LARGE TABLE ISOLATORS

MICRO-g pedestal isolators are the answer to isolating large research tables. If you have an existing top, tell us the size and weight and we'll provide the correct isolators. Or, we can supply everything, including a granite plate or structured steel top, at competitive prices. Standard isolator components will handle loads from a few hundred pounds to several tons.



ADVANCED, GIMBAL PISTON* WITH HORIZONTAL ISOLATION

The thin-wall diaphragm isolator element, such as we use, has always been best for vertical isolation. Now, we've developed a way to make it virtually as effective in the horizontal mode as the vertical. It's a new type isolator called the TMC Gimbal Piston*. It handles horizontal inputs with an effectiveness derived from its superb vertical characteristics. It is entirely stable, with no sideways lean tendencies, and does not depend on the length of a hanging rod, wire or counter-mass for its frequency characteristics. The TMC Gimbal Piston is optional on all sizes of MICRO-g isolators. It is simply the best isolator we've ever measured.

*Patent Applied For

WRITE OR CALL FOR LITERATURE OR PROPOSAL

TECHNICAL MANUFACTURING CORPORATION

PRECISION METAL PRODUCTS, VIBRATION ISOLATION SYSTEMS
185 NEW BOSTON STREET, WOBURN, MA 01801 • 617-933-0050

Circle No. 16 on Reader Enquiry Card.



LETTERS

R.M. Tucker and M.E. Tucker	Evidence of synsedimentary tectonic movements in the Triassic halite of Cheshire	495	A. McLaren, P. Chandler, M. Buehr, W. Fierz and E. Simpson	Immune reactivity of progeny of tetraparental male mice	513
J. Charteris, J.C. Wall and J.W. Nottrodt	Functional reconstruction of gait from the Pliocene hominid footprints at Laetoli, northern Tanzania	496	W. Hunkeler, H. Möhler, L. Pieri, P. Polc, E.P. Bonetti, R. Cumin, R. Schaffner and W. Haefely	Selective antagonists of benzodiazepines	514
J.E. Lloyd	Firefly mate-rivals mimic their predators and vice versa	498	M. Chinkers and S. Cohen	Purified EGF receptor-kinase interacts specifically with antibodies to Rous sarcoma virus transforming protein	516
R. Lubbock and W.B. Amos	Removal of bound calcium from nematocyst contents causes discharge	500	J.E. Kudlow, J.E. Buss and G.N. Gill	Anti-pp60^{src} antibodies are substrates for EGF-stimulated protein kinase	519
D.S. Secher	Immunoradiometric assay of human leukocyte interferon using monoclonal antibody	501	B.R. Jordan, F. Bregegere and P. Kourilsky	Human HLA gene segment isolated by hybridization with mouse H-2 cDNA probes	521
M. Shapiro and R.P. Erickson	Evidence that the serological determinant of H-Y antigen is carbohydrate	503	H.I. Miller and H.A. Nash	Direct role of the <i>himA</i> gene product in phage λ integration	523
R.A.G. Smith, R.J. Dupe, P.D. English and J. Green	Fibrinolysis with acyl-enzymes: a new approach to thrombolytic therapy	505	R.Y. Tsien	A non-disruptive technique for loading calcium buffers and indicators into cells	527
L. Brent, L.S. Rayfield, P. Chandler, W. Fierz, P.B. Medawar and E. Simpson	Supposed lamarckian inheritance of immunological tolerance	508			

BOOK REVIEWS

Rudolf Peierls	Reminiscences of Los Alamos, 1943 – 1945 (edited by L. Badash <i>et al.</i>)	529	G. Gee	Crystal Melting (by B. Wunderlich)	531
Peter D. Moore	Land, Man, and Sand (by J. Walls)	529	Henry Charnock	The North-West European Shelf Seas, Vol. II (edited by F.T. Banner <i>et al.</i>)	532
Lindon J. Eaves	The IQ Game (by H.F. Taylor)	530	E. Arthur Bell	The Biosynthesis of Mycotoxins (edited by P.S. Steyn)	532
Ben Mifflin	Microorganisms and Nitrogen Sources (edited by J.W. Payne)	531			

MISCELLANY

100 years age	448
Classified advertising	xliv

PRODUCT REVIEW This week's product review on p. xxxv covers products on view at the FASEB exhibition.

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.
 - Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.
 - Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.
 - 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.
- Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.
- Manuscripts** may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: exp (a) is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov. (1976).

New from Nuclepore

Ultrafiltration membranes

Circle No. 19 on Reader Enquiry Card.

- Lowest adsorption losses
- Highest chemical resistance, with autoclavability
- Widest range of molecular weight cut-offs (50-1,000,000)
- High flow rates

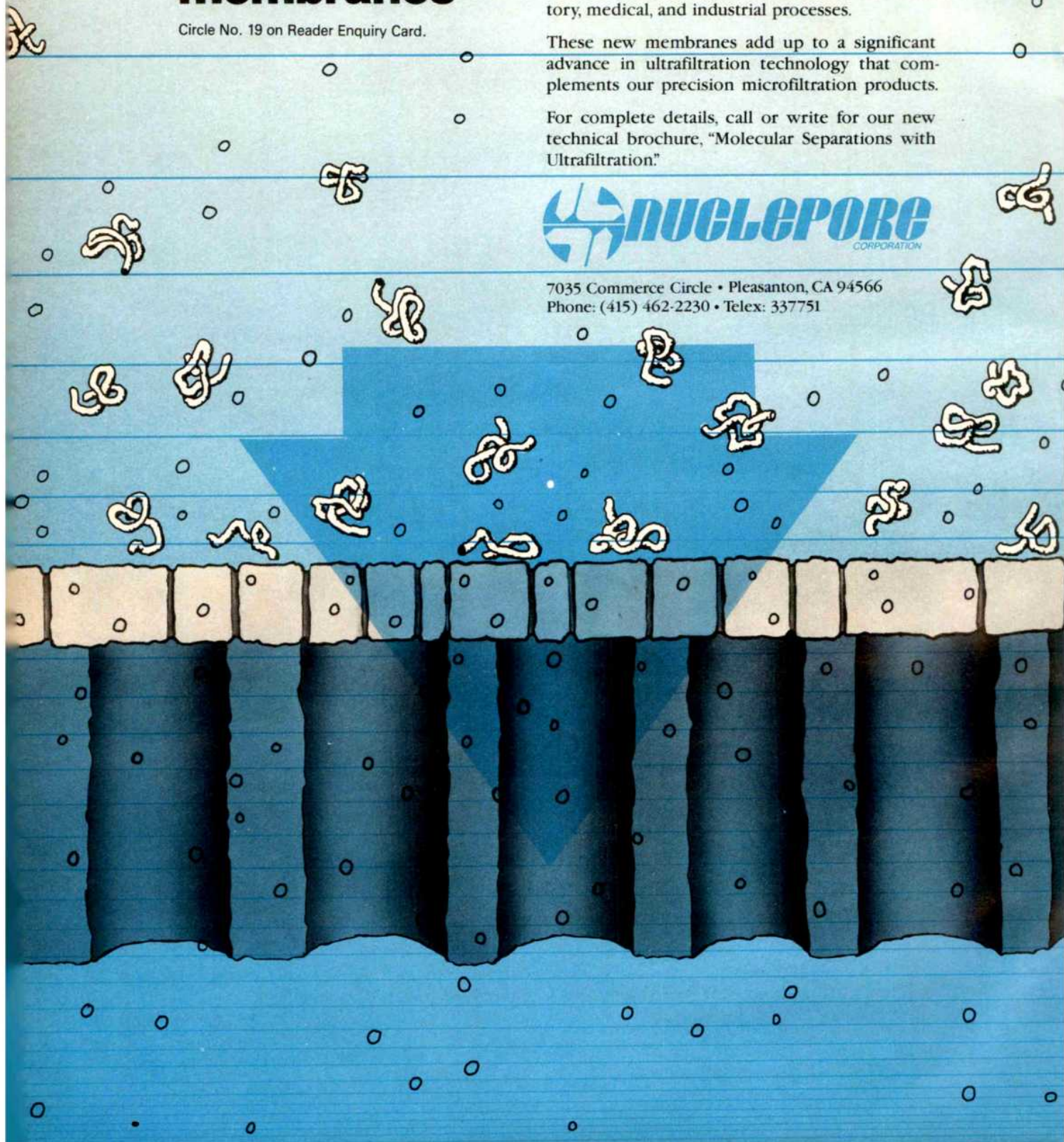
Nuclepore, the membrane innovation company of the 80's, has introduced three new ultrafiltration membranes for molecular separations that offer new levels of precision to research, laboratory, medical, and industrial processes.

These new membranes add up to a significant advance in ultrafiltration technology that complements our precision microfiltration products.

For complete details, call or write for our new technical brochure, "Molecular Separations with Ultrafiltration."



7035 Commerce Circle • Pleasanton, CA 94566
Phone: (415) 462-2230 • Telex: 337751



nature

Volume 290 No.5806 9 April 1981

Per Ardua ad Astra (et Terram)	433	Nature misrepresented	437
Who, not Congress, should police fraud?	433	Known misdemeanours	438
Universities complain at Pentagon policy: Restrictions on foreign students cause alarm	435	UK hydraulics research: Deep water ahead	438
Gowans washes hands of Rothschild	435	Europe's directive on drugs: Lords back industry	439
DNA research guidelines: Dutch get tough	436	Tangled policy on drugs	439
Solar polar mission: Joint project saved?	436	Opening up Siberia: Research in steppes	439
Congress redispenses	436	Correspondence: Cancer hazards (R.A. Stallones and T.D. Downs); Draize test must go (L. Sargeant); IQ put to test (T. Nettelbeck and M. Lally); Why conserve? (M.F. Heyworth and J.S. Wainscoat); Lamarck's revival (V. Klemes); Evolution on trial (T.H. Jukes)	440
Biotechnology: Canada takes stock	436		
European uranium supply: Australia signs up	437		
Fraud in research: Congress asks why	437		

NEWS AND VIEWS

A tropical Volute shell and the Icarus syndrome: Jonathan C. Howard on the failure to confirm reports of inheritance of immunological tolerance	441	Stable clocks and the theory of relativity: G.J. Edwards on an experimental confirmation of the change in the rate of a clock moving in a gravitational field	446
Small is beautiful: Piet Borst and Leslie A. Grivell compare the human and yeast mitochondrial genome	443	Quasars — what and where?: Martin Ward reports a recent meeting	447
Thrombosis: P.J. Gaffney looks at a new molecular approach to thrombolytic therapy	445	Triggering of volcanic eruptions: R.S.J. Sparks comments on the role of Earth tides	448

REVIEW ARTICLE

M. Klingenberg

Membrane protein oligomeric structure and transport function 449

ARTICLES

S.F. Dermott	The 'braided' F-ring of Saturn	454	J. Montoya, D. Ojala and G. Attardi	Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs	465
S. Anderson, A.T. Bankier, B.G. Barrell, M.H.L. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J.H. Smith, R. Staden and I.G. Young	Sequence and organization of the human mitochondrial genome	457	D. Ojala, J. Montoya and G. Attardi	tRNA punctuation model of RNA processing in human mitochondria	470
			W.S. Hayward, B.G. Neel and S.M. Astrin	Activation of a cellular <i>onc</i> gene	475

LETTERS

B. Margon, G.A. Chanan and R.A. Downes	A companion quasar to 3C345	480	B.G. Ferguson	Measurement of ionospheric winter anomaly in Southern Hemisphere	485
M.G. Edmunds and G.H. George	Random alignment of quasars	481	R.H. Byrne	Inorganic lead complexation in natural seawater determined by UV spectroscopy	487
E. Boldt and D. Leiter	Precursor active galaxies and the cosmic X-ray background	483	J.L. LaBrecque and P. Barker	The age of the Weddell Basin	489
			M.J.C. Nutt and E.G. Smith	Transcurrent faulting and the anomalous position of pre-Carboniferous Anglesey	492

Cover: Interference micrograph of nematocysts of *Rhodactis rhodostoma* in various stages of discharge. Changes in the refractive index of the capsule contents reflect dilution of the capsular fluid. A new theory for nematocyst discharge is described on p.500.

The sequence of human mitochondrial DNA is reported on p.457 and this paper is discussed on p.443 in *News and Views*.

M

Nature (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd.

Contents continued overleaf

Definitive, Complete, Authoritative

NATURE announces the publication in 1981 of the Annual Directory of Biologicals and Disposable Chemicals. The directory will be available for distribution to all NATURE subscribers in the Autumn of 1981. It will also be marketed separately, offering an even larger distribution.

A New Buyers guide

The basic thrust of the directory will be to provide NATURE subscribers with a definitive reference catalogue from which they may purchase the chemicals and disposables listed.

From research conducted around the world the NATURE marketing staff has determined the need for a directory of biologicals, disposable chemicals and gases. Probably more than any other scientist, the NATURE reader represents the largest potential market for the sale of these chemicals.

Every attempt will be

made to produce a catalogue that will reference the worldwide market of chemical manufacturers, distributors, suppliers, and exporters.

No instrumentation or apparatus listings will be included in this directory; its sole purpose is to provide detailed information for our readers as to where they may secure the biologicals, chemicals and gases for use in their laboratories.

The directory will be cross-referenced and include chemical classifications, product classifications, geographical location of suppliers and other referenced information as needed.

NATURE Directory of Biologicals



What's in It?

The directory will include (but not be limited to) these biological research chemicals

- immunochemicals
 - cell lines
- fine chemicals
 - reagents
 - enzymes
 - isotopes
- monoclonals
- gases, etc.

Companies write for details NOW!

I'm interested in knowing more about Nature's new Directory of Biologicals.

☐ Send me a questionnaire so that my company can have its products listed in the Directory

☐ I am interested in reserving advertising space, send me a rate card for the Directory

☐ Put me on your mailing list to receive more information

☐ Have a representative call or see me

My name: _____

Title: _____

Company: _____

Address: _____

Zip/Code _____

Telephone number: _____

nature

Circle No.08 on Reader Enquiry Card.

**Andy Sutherland
or Marion Delaney
Nature
4 Little Essex Street
London WC2R 3LF**

**Henry Dale
Nature
15 East 26 Street
New York NY 10010**

LETTERS

- | | | | | | |
|---|---|------------|---|--|------------|
| R.S. Harmon, R.S. Thorpe
and P.W. Francis | Petrogenesis of Andean andesites
from combined O-Sr isotope
relationships | 396 | K. Umezawa, I.B. Weinstein,
A. Horowitz, H. Fujiki,
T. Matsushima and T. Sugimura | Similarity of teleocidin B and
phorbol ester tumour promoters in
effects on membrane receptors | 411 |
| R. Meissner, H. Bartelsen
and H. Murawski | Thin-skinned tectonics in
the northern Rhenish
Massif, Germany | 399 | A. Baranyi and O. Fehér | Synaptic facilitation requires
paired activation of convergent
pathways in the neocortex | 413 |
| R.K. Colwell | Group selection is implicated in
the evolution of female-biased
sex ratios | 401 | R. Payne and J. Howard | Response of an insect
photoreceptor: a simple log-
normal model | 415 |
| U.P. Devaskar, S.U. Devaskar,
S. Voina, N. Velayo
and M.A. Sperling | Growth hormone stimulates adrenal
steroidogenesis in the fetus | 404 | J.K. Blusztajn
and R.J. Wurtman | Choline biosynthesis by a
preparation enriched in
synaptosomes from rat brain | 417 |
| A.W. Segal, M. Geisow,
R. Garcia, A. Harper
and R. Miller | The respiratory burst of phagocytic
cells is associated with a rise
in vacuolar pH | 406 | R.W. Voellmy
and A.L. Goldberg | ATP-stimulated endoprotease is
associated with the cell membrane
of <i>E. coli</i> | 419 |
| K.S. Landreth, K. McCoy, J.
Clagett, F.J. Bollum
and C. Rosse | Deficiency in cells expressing
terminal transferase in
autoimmune (motheaten) mice | 409 | J.W. Roberts
and C.W. Roberts | Two mutations that alter the
regulatory activity of <i>E. coli</i>
recA protein | 422 |
| | | | S. Hill, C. Kennedy,
E. Kavanagh, R.B. Goldberg
and R. Hanau | Nitrogen fixation gene (<i>nifL</i>)
involved in oxygen regulation of
nitrogenase synthesis in
<i>K. pneumoniae</i> | 424 |

MATTERS ARISING

- | | | | | | |
|---------------------------------|---|------------|-------------------------------|---|------------|
| S.W. Hui and T.P. Stewart | 'Lipidic particles' are inter-
membrane attachment sites | 427 | F.R.S. Clark and D.A. Russell | Fossil charcoal and the
palaeoatmosphere | 428 |
| A.J. Verkleij and B. de Kruijff | Reply | 427 | M.J. Cope and W.G. Chaloner | Reply | 428 |

BOOK REVIEWS

- | | | | | | |
|------------------|--|------------|--------------|--|------------|
| Eric Ashby | Implementing Public Policy
(by Ruth Levitt) | 429 | C.M. Yonge | The Functioning of Freshwater
Ecosystems (edited by
E.D. Le Cren and
R.H. Lowe-McConnell) | 431 |
| Jifi Novotný | Evolution and Variation of
Multigene Families
(by Tomoko Ohta) | 430 | F.W. Shotton | The Shaping of Southern England
(edited by David K.C. Jones) | 431 |
| E. Philip Krider | Ball Lightning and Bead
Lightning: Extreme Forms of
Atmospheric Electricity (by
James Dale Barry) | 430 | | | |

MISCELLANY

- | | | | |
|--------------------------|------------|-------------------------------|------------|
| 100 years ago | 362 | Classified advertising | xii |
| New on the market | xi | | |

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.
 - Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.
 - Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.
 - 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.
- Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.
- Manuscripts** may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in **258**, 1 (1975) and **264**, v, 11 Nov. (1976).

ANNOUNCING A MAJOR NEW SERIES FROM MACMILLAN PUBLISHERS

TOPICS IN MOLECULAR & STRUCTURAL BIOLOGY

General Editors:

WATSON FULLER,

Department of Physics, University of Keele.

STEPHEN NEIDLE,

Department of Biophysics, King's College, London.

TOPICS IN MOLECULAR AND STRUCTURAL BIOLOGY is a major new series which is aimed at emphasising interest in three-dimensional structure – the feature which most obviously characterises molecular biology as a discipline.

Each volume will focus on a particular family of biological macromolecules or a well-defined area of structural interest. The contributions will be chosen to reflect the wide range of physical and chemical techniques which characterise contemporary approaches to problems in molecular biology.

The aim will be to highlight topics of high current interest rather than to provide a text with uniform and comprehensive treatment of the field.

The contribution that this series will make in encouraging interest in three-dimensional structure will be clearly derived from a broad experimental base. Descriptions of more speculative model building and conformational analysis will certainly not be excluded but there will be a major emphasis on identifying the assumptions which limit their significance.

The need to place such studies in some perspective would seem to be particularly timely since the rapidly increasing availability within many laboratories of high speed computers with good quality graphics can be expected to stimulate a great deal of work in this area.

The general level of treatment in this series is aimed at research workers in molecular biology and related fields. The major emphasis is on the results which have been obtained and the significance which can be attached to them, rather than on details of the techniques which have been used. This is an approach which should make these volumes of particular value to workers in cognate areas of applied biology and medicine.

VOLUME 1: TOPICS IN NUCLEIC ACID STRUCTURE –

Edited by STEPHEN NEIDLE

Spring 1981; £20; c230pp; ISBN 0 333 26678 1

This work systematically details the most recent data on all structural aspects of oligonucleotides, transfer RNAs, polymeric nucleic acids and chromatin.

Each category is described both in terms of X-ray crystallographic and solution NMR data. A standardised and rational system of nomenclature is used throughout the book.

VOLUME 2: FURTHER TOPICS IN NUCLEIC ACID STRUCTURE –

Winter 1981 c£24 ISBN 0 333 31555 3

VOLUME 3: MOLECULAR ASPECTS OF ANTICANCER DRUG ACTION –

Winter 1981 c£24. ISBN 0 333 31556 1

Send your orders, requests for further information etc. to: Frances Roach, Product Manager, Scientific & Medical Division, Macmillan Publishers Ltd, Houndmills, Basingstoke, Hampshire RG21 2XS, U.K.

M

NEWS

Personal privacy is absolute privacy	349	UK computer industry: ICL falters	353
Scenarios on energy	349	London medical schools: Mergers galore	354
Congress writes back Reagan's economies: Democrats prepare for battle	351	Law of the Sea: US backtracks	354
Hidden cheer culled from statistics	351	India mines the sea	355
UK aerospace: Eager for orders	352	Correspondence: Licking Darwin (G.P. Vinson); Weed control (N.W. Pirie); Cafeteria feeding (A.S. Cole); Cancer causation (J.E. Trosko); Genes and free will (S. Rose); Conservation sites (R.H.L. Disney); DMSO and immunity (M.J. Ashwood-Smith; A. Pestronk and D.B. Drachman)	356, 432
UK nuclear power: Battle begun	352		
Czech nuclear power: Action on safety	353		
Soviet space programme: Mongolia's turn	353		

NEWS AND VIEWS

Nerves, myelin and multiple sclerosis: Miranda Robertson on a meeting of clinicians with research biologists	357	The origin of cosmic gamma-ray bursts — new Soviet results: B.J. Teegarden surveys recent results and resulting hypotheses	361
A new model for phototransduction: Peter Lillywhite assesses an alternative to the cascade model	358	Plasmids and bacterial pathogens: J.R. Saunders reports a meeting held in Santo Domingo	362
Second sound in ^3He: P.V.E. McClintock discusses propagating waves in superfluid ^3He	359	Superluminal radio sources: Guy Pooley explains why we should believe in them	363
Ice core studies: dating the past to find the future: W. Dansgaard stresses the importance of dating ice cores	360	DNA methylation and control of gene expression: Tomas Lindahl looks for roles for methylated bases in DNA	363

ARTICLES

T.J. Pearson, S.C. Unwin, M.H. Cohen, R.P. Linfield, A.C.S. Readhead, G.A. Seielstad, R.S. Simon and R.C. Walker	Superluminal expansion of quasar 3C273	365	C. Coleclough, R.P. Perry, K. Karjalainen and M. Weigert	Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression	372
P.A. Hieter, S.J. Korsmeyer, T.A. Waldmann and P. Leder	Human immunoglobulin λ light-chain genes are deleted or rearranged in λ-producing B cells	368			

LETTERS

E.P. Mazets, S.V. Golenetskii, R.L. Aptekar', Yu.A. Gur'yan and V.N. Il'inskii	Cyclotron and annihilation lines in γ-ray bursts	378	D.J. Barber, L.A. Freeman and D.J. Smith	High-voltage, high-resolution lattice images of dolomite	389
R.S. Booth, A.J. Kus, R.P. Norris and N.D. Porter	Observations of a circumstellar shell around the OH/IR star OH127.8-0.0	382	T.M. Dillon, J.G. Richman, C.G. Hansen and M.D. Pearson	Near-surface turbulence measurements in a lake	390
M.A. Hapgood and P. Rothwell	Fragmentation of a meteor in near-Earth space	384	L.V. Medford, A. Meloni, L.J. Lanzerotti and G.P. Gregori	Geomagnetic induction of a transatlantic communications cable	392
R. Sinclair, T. Yamashita and F.A. Ponce	Atomic motion on the surface of a cadmium telluride single crystal	386	D.V. Kent, D. Ninkovich, T. Pescatore and S.R.J. Sparks	Palaeomagnetic determination of emplacement temperature of Vesuvius AD 79 pyroclastic deposits	393

Contents continued overleaf

Cover: High-resolution radio astronomy observations of the quasar 3C273 over a period of three years provide the first direct evidence of faster-than-light expansion in a quasar. Detailed radio maps and results are presented on page 365. On page 363 Guy Pooley discusses faster-than-light radio sources in general.

M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd.

Radiopharmaceuticals

for
diagnostic and
therapeutic
applications

high radiochemical purity
sterilization according to DAB 7

comprehensive range
favourable delivery terms



isocommerz

GmbH · Binnen- und Außenhandelsunternehmen
für radioaktive und stabile Isotope
DDR – 1115 Berlin, Lindenberger Weg 70
German Democratic Republic

Circle No. 20 on Reader Enquiry Card.

LETTERS

D.B. Prior, Wm J. Wiseman Jr and W.R. Bryant	Submarine chutes on the slopes of fjord deltas	326	W.H.M. Peters, H.G.P. Swarts, J.J.H.H.M. de Pont, F.M.A.H. Schuurmans Stekhoven and S.L. Bonting	(Na⁺ + K⁺)ATPase has one functioning phosphorylation site per α subunit	338
S.J. Culver and M.A. Buzas	Foraminifera distribution of provinces in the Gulf of Mexico	328	L.B. Vogler, J.L. Preud'homme, M. Seligmann, W.E. Gathings W.M. Crist, M.D. Cooper and F.J. Bollum	Diversity of immunoglobulin expression in leukaemic cells resembling B-lymphocyte precursors	339
R.V. Bensasson, E.J. Land, A.L. Moore, R.L. Crouch, G. Dirks, T.A. Moore and D. Gust	Mimicry of antenna and photoprotective carotenoid functions by a synthetic carotenoporphyrin	329	C. McLean and P.J. Lowry	Natural occurrence but lack of melanotrophic activity of γ-MSH in fish	341
N.V. Swindale	Absence of ocular dominance patches in dark-reared cats	332	C.J. Hanna, M.K. Bach, P.D. Pare and R.R. Schellenberg	Slow-reacting substances (leukotrienes) contract human airway and pulmonary vascular smooth muscle <i>in vitro</i>	343
B.W. Matthews, M.G. Grütter, W.F. Anderson and S.J. Remington	Common precursor of lysozymes of hen egg-white and bacteriophage T4	334			
J.H. Coggin Jr, J.E. Oakes, R.J. Huebner and R. Gilden	Unusual filterable oncogenic agent isolated from horizontally transmitted Syrian hamster lymphomas	336			

BOOK REVIEWS

Richard Dawkins	Programmed to Learn: An Essay on the Evolution of Culture (by H. Ronald Pulliam and Christopher Dunford)	345	Walter A. Harrison	Electrons at the Fermi Surface (edited by M. Springford)	347
C.W. Hilbers	Nucleic Acid Geometry and Dynamics (edited by R.H. Sarma)	346	Peter D. Moore	The Northwest European Pollen Flora, Vol. II (edited by W. Punt and G.C.S. Clarke)	348
N. Sharon	Bacterial Adherence (edited by E.H. Beachey)	346	I.S. Grant	Deep-Inelastic and Fusion Reactions with Heavy Ions (edited by W. von Oertzen)	348

MISCELLANY

100 years ago	292
Author Index to Vol.289	xiii

Classified advertising

Pat. Ch. City Lib. xxvi
SCIENCE COLLEGE
A Special Group for
REFERENCE

GUIDE TO AUTHORS

● Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.

● Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.

● Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.

● 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

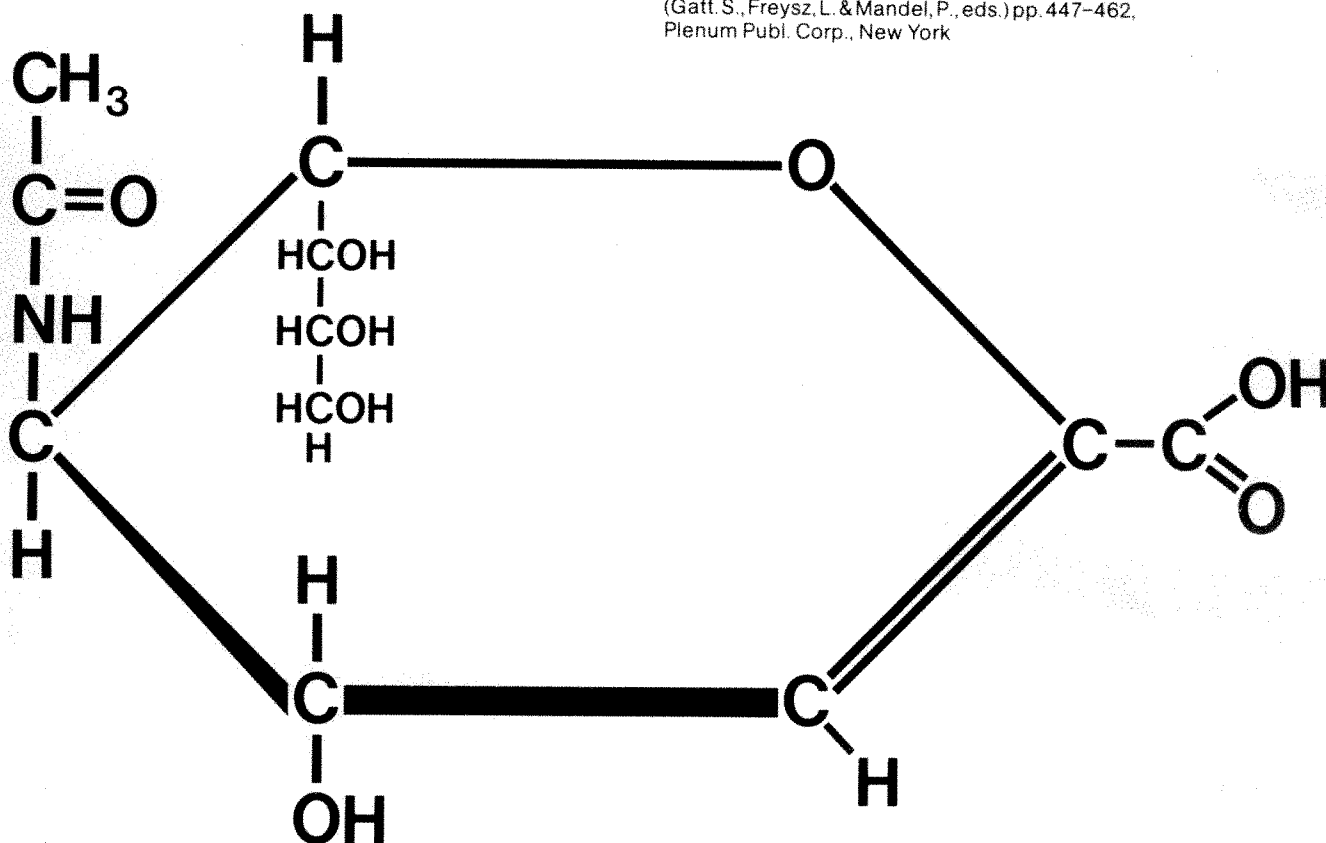
In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov. (1976).

2,3-Dehydro-2-deoxy-N-acetylneuraminic acid

the competitive inhibitor of neuraminidases

Veh, R. W. & Schauer, R. (1978) in *Enzymes of Lipid Metabolism* (Gatt, S., Freysz, L. & Mandel, P., eds.) pp. 447-462, Plenum Publ. Corp., New York



is now available with
BOEHRINGER MANNHEIM
 your supplier of:

N-acetylneuraminic acid	Cat. No. 101 923
N-acetylneuraminosyl-D-lactose	Cat. No. 161 721
N-acetylneuraminic acid glycoprotein	Cat. No. 102 032
2,3-dehydro-2-deoxy-N-acetylneuraminic acid	Cat. No. 528 544
2-(3'-methoxyphenyl)-N-acetyl-D-neuraminic acid	Cat. No. 174 696
neuraminic acid-β-methyl glycoside	Cat. No. 107 603
neuraminidase from <i>Arthrobacter ureafaciens</i>	Cat. No. 269 611
neuraminidase from <i>Clostridium perfringens</i>	Cat. No. 107 590

For details please contact your local representative or



Boehringer Mannheim GmbH
 Biochemica
 P.O. Box 310120
 D-6800 Mannheim
 W. Germany

nature

Volume 290 No.5804 26 March 1981

NEWS

Reagan's non-existent science advisors	279	UK astronomy: Change of direction	283
How not to run telecommunications	279	Oak tree wilt: Europe in dread	284
DNA guidelines for further attenuation: NIH to consider new voluntary provisions	281	Soviet higher education: Candidates change	284
Guidelines transgressor let off lightly	281	Irish research and development: Marching to Georgia	284
UK research councils: Spending too soon	282	Bulgarian astronomers look up and abroad	285
Polish academics: Degrees in question	282	UK universities: Jobs to be cut?	285
US universities: Commercial interest	283	Correspondence: Toxicity testing (L. Sargeant); Inherited tolerance (K. W. Jones); Exhibit dismay (G. G. Simpson); Evolving theory (R. V. Hesketh; P. L. Miller)	286
New NASA man	283		

NEWS AND VIEWS

Genes pieced together: P.J. Artymiuk, C.C.F. Blake and A.E. Sippel find exonic regions of hen lysozyme in phage lysozyme	287	Hybrid myosins: Dixie W. Frederiksen explains new insights into the control of muscle contraction	291
The American Meteorological Society's January meeting: Waves, turbulence, solar convection and ozone: Michael McIntyre	288	Ocean plateaus: Peter J. Smith on the evidence for two distinct types	292
Climatic variations: M.J. Salinger and C.B. Sear	289	Idiotypic restriction in myasthenia gravis antibodies: Angela Vincent reports a recent meeting	293
The return of Halley: David W. Hughes describes the early signs of the comet's return	290	Disposing of dioxins: Alistair Hay explains the potential of oxidation	294

ARTICLES

G. Kukla, A. Berger, R. Lotti and J. Brown	Orbital signatures of interglacials	295	C. Benoist and P. Chambon	<i>In vivo</i> sequence requirements of the SV40 early promoter region	304
J.J. Walsh	A carbon budget for overfishing off Peru	300	D.J. Mathis and P. Chambon	The SV40 early region TATA box is required for accurate <i>in vitro</i> initiation of transcription	310

LETTERS

A.K. Harding and F.W. Stecker	Pulsar and diffuse contributions to observed galactic γ radiation	316	R.G. Ross, P. Andersson and G. Bäckström	Unusual <i>PT</i> dependence of thermal conductivity for a clathrate hydrate	322
R.T. Schilizzi, G.K. Miley, J.D. Romney and R.E. Spencer	Changes in orientation and appearance of the 100 m arcs radio structure in SS433	318	D.C. Ayres	Destruction of polychlorodibenzo-<i>p</i>-dioxins	323
H. Böhringer and F. Arnold	Acetonitrile in the stratosphere — implications from laboratory studies	321	J.R. Wilson, K. H. Esbensen and P. Thy	A new pyroxene fractionation trend from a layered basic intrusion	325

Contents continued overleaf

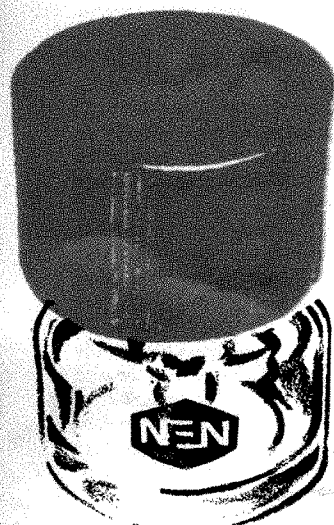
Cover: Photograph of an artist's drawing of a species of *Elphidium* from Florida. S.J. Culver and M.A. Buzas describe on page 328 the distribution of such benthic foraminifera in the Gulf of Mexico.

M

Nature (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd.

74 Research Assistants for

Receptor Site Studies



Circle No.06 on Reader Enquiry Card.

α -Adrenergic

Clonidine hydrochloride, [4-³H]-
Desmethylinipramine hydrochloride,
[2,4,6,8-³H]-
Dihydro- α -ergocryptine, 9,10-[9,10-³H(N)]-
WB-4101 (2,6-Dimethoxyphenoxyethyl)
aminomethyl-1, 4-benzodioxane,
2-[*phenoxy*-3-³H(N)]-
Epinephrine, *levo*-[*methyl*-³H]-
Norepinephrine, *levo*-[7,8-³H(N)]-

β -Adrenergic

Carazolol, DL-[3,6-³H(N)]-
Dihydroalprenolol hydrochloride, *levo*-
[*propyl*-2,3-³H]-
Epinephrine, *levo*-[*N-methyl*-³H]-
Hydroxybenzylisoproterenol, *p*-[7-³H]-
Iodohydroxybenzylpindolol, [1²⁵]-
Isoproterenol, DL-[7-³H(N)]-
Norepinephrine, *levo*-[7,8-³H(N)]-
Propranolol, L-[4-³H]-

Aspartate

Aspartic acid, D-[2,3-³H]-
Aspartic acid, L-[2,3-³H]-
Methyl-D-aspartic acid, *N*-[*methyl*-³H]-

Benzodiazepine

Diazepam, [*methyl*-³H]-
Flunitrazepam, [*methyl*-³H]-

Cholinergic

Muscarinic

Acetylcholine chloride, [*N-methyl*-³H]-
Choline chloride, [*methyl*-³H]-
Pilocarpine, [³H(G)]-
Quinuclidinyl benzilate,
DL-[*benzyl*-4,4'-³H(N)]-
Scopolamine methyl chloride,
[*N-methyl*-³H]-

Nicotinic

Acetylcholine chloride, [*N-methyl*-³H]-
 α -Bungarotoxin, [1²⁵]-
Choline chloride, [*methyl*-³H]-
Tubocurarine chloride, *dextro*-[13'-³H(N)]-

Dopaminergic

ADTN Amino-6,7-dihydroxy-
1,2,3,4-tetrahydronaphthalene,
2-[5,8-³H]-
Amphetamine sulfate, D-[³H(G)]-
Apomorphine, [8,9-³H]-
Chlorpromazine, [³H]-
Dihydroxyphenylethylamine,
3,4-[*ethyl*-1-³H(N)]- or [*ethyl*-2-³H(N)]-
Haloperidol, [³H(G)]-
Propylnorapomorphine, *N*-[*propyl*-³H(N)]-
Spiroperidol, [1-*phenyl*-4-³H]-

GABA

Alanine, β -[3-³H(N)]-
Aminobutyric acid, γ -[2,3-³H(N)]-
Dihydropicrotoxinin, α -[8,10-³H]-
Isoguvacine hydrochloride, [³H]-
Muscimol, [*methylene*-³H(N)]- or [4-³H]-
Nipecotic acid, [*ring*-³H]-

Glutamate

Glutamic acid, L-[3,4-³H]-

Glycine

Glycine, [2-³H]-

Histamine

H₁

Histamine, [³H(G)]-
Pyrilamine, [*pyridinyl*-5-³H]- (Mepyramine)

H₂

Histamine, [³H(G)]-

Opiate

Dihydromorphine, [7,8-³H(N)]-
Enkephalin (5-L-leucine), [*tyrosyl*-3,5-³H(N)]-
Enkephalin (5-L-methionine),
[*tyrosyl*-3,5-³H(N)]-
Enkephalinamide
(2-D-alanine-5-L-methionine),
[*tyrosyl-ring*-2,6-³H]-
Ethylketocyclazocine, [9-³H]-
Morphine, [6-³H(N)]-

Serotonin

Hydroxytryptamine binoxalate, 5-[1,2-³H(N)]-
Hydroxytryptamine creatinine sulfate,
5-[1,2-³H(N)]-

Steroid

Androgen

Dihydrotestosterone,
[1,2,4,5,6,7,16,17-³H(N)]-
Methyltrienolone, [17 α -*methyl*-³H]- (R1881)*
Testosterone, [1,2,6,7,16,17-³H(N)]-

Estrogen

Estradiol, [2,4,6,7,16,17-³H(N)]-
Iodo-3, 17 β -estradiol, 16 α -[1²⁵]-
Moxestrol, [11 β -*methoxy*-³H]- (R2858)*

Glucocorticoid

Dexamethasone, [6,7-³H(N)]-
Prednisolone, [6,7-³H(N)]-
Triamcinolone acetonide, [6,7-³H(N)]-

Mineralocorticoid

Aldosterone, D-[1,2,6,7-³H(N)]-

Progesterone

Dihydroprogesterone, [1,2-³H(N)]-
Nor-17 α -ethynyltestosterone, 19-[6,7-³H(N)]-
Progesterone, [1,2,6,7-³H(N)]-
Promegestone, [17 α -*methyl*-³H]- (R5020)*

*Manufactured by NEN under licensed agreement of
ROUSSEL-UCLAF.

Miscellaneous

Dihydroxyvitamin D₃, 1 α , 25-[26,27-³H]-
Hydroxyvitamin D₃, 25-[26,27-³H]-
Imipramine hydrochloride, [2,4,6,8-³H]-
Melanocyte stimulating hormone inhibiting
factor, [*proline*-2,3,4,5-³H]- (MIF)
Phencyclidine, [*piperidyl*-3,4-³H(N)]-
Reserpine, [*benzoyl*-³H(G)]-



New England Nuclear

549 Albany Street, Boston, Massachusetts 02118
Call toll-free 800-225-1572
(In Massachusetts and International: 617-482-9595)

NEN Chemicals GmbH: D-6072 Dreieich, W. Germany
Postfach 401240 Telephone (06103) 85034 Telex 4-17993 NEN D
NEN Canada Ltd., 2453 46th Avenue, Lachine, Que H8T 3C9
Telephone 514-636-4971 Telex 05-821808

A.G.J. Rhodin, J.A. Ogden and G.J. Conlogue	Chondro-osseous morphology of <i>Dermochelys coriacea</i>, a marine reptile with mammalian skeletal features	244	G.T. Pearson, J.S. Davison, R.C. Collins and O.H. Petersen	Control of enzyme secretion by non-cholinergic, non-adrenergic nerves in guinea pig pancreas	259
A.R. Blaustein and R.K. O'Hara	Genetic control for sibling recognition?	246	C. Shih, L.C. Padhy, M. Murray and R.A. Weinberg	Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts	261
J. Hochman, A. Katz, E. Levy and S. Eshel	Substrate-adhering lymphoid cells show impaired tumorigenicity and increased immunogenicity	248	A.R. Stuitje, C.E. Spelt, E. Veltkamp and H.J.J. Nijkamp	Identification of mutations affecting replication control of plasmid Clo DF13	264
A.K. Harris, D. Stopak and P. Wild	Fibroblast traction as a mechanism for collagen morphogenesis	249	D. Hentzen, A. Chevallier and J.-P. Garel	Differential usage of iso-accepting tRNA^{Ser} species in silk glands of <i>Bombyx mori</i>	267
E.M. Eugui and D.L. Emery	Genetically restricted cell-mediated cytotoxicity in cattle immune to <i>Theileria parva</i>	251	P. Downes and R.H. Michell	Human erythrocyte membranes exhibit a cooperative calmodulin-dependent Ca²⁺-ATPase of high calcium sensitivity	270
O. Stutman and M.J. Cuttito	Normal levels of natural cytotoxic cells against solid tumours in NK-deficient beige mice	254	H. Takisawa and M. Makinose	Occluded bound calcium on the phosphorylated sarcoplasmic transport ATPase	271
M. Reth, G. Kelsoe and K. Rajewsky	Idiotypic regulation by isologous monoclonal anti-idiotope antibodies	257	A. Gronenborn, B. Birdsall, E.I. Hyde, G.C.K. Roberts, J. Feeney and A.S.V. Burgen	Direct observation by NMR of two coexisting conformations of an enzyme-ligand complex in solution	273

Part of the Library
SCIENCE COLLEGE
4 Upper Circular Road
SINGAPORE

BOOK REVIEWS

S.J. Goldsack	A History of Computing in the Twentieth Century (edited by N. Metropolis <i>et al.</i>). Early British Computers (by S. Lavington)	275	Michael Bulmer	Evolution in Age-Structured Populations (by Brian Charlesworth)	277
P.W. Hawkes	The Computer in Optical Research (edited by B.R. Frieden)	276	A. Serafini-Fracassini	Biology and Pathology of Elastic Tissues (edited by A.M. Robert and L. Robert)	278
Andrew Hamnett	The Double Layer (edited by J. O'M. Bockris <i>et al.</i>)	276	R.J. Hance	The Bipyridinium Herbicides (by L.A. Summers)	278

MISCELLANY

Product review — Laboratory computers	xxiii
Classified advertising	xxv
100 years ago	192

GUIDE TO AUTHORS

- Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.
- Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.
- Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.
- 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*. Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

- only one reference number need be used if the reference is to several papers by identical authors.
- first and last pages of references should be cited.

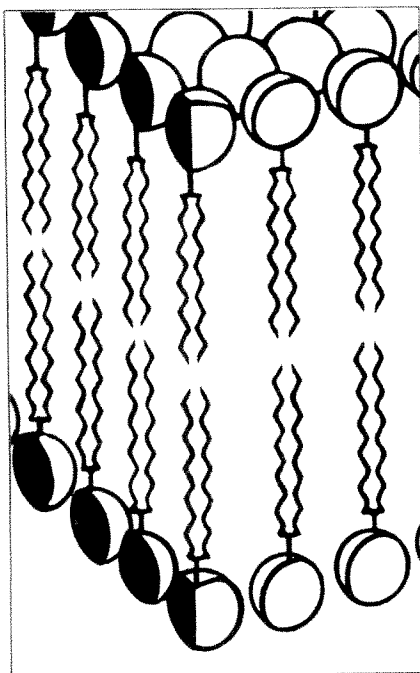
Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov. (1976).

¹⁴C P o-... oli-i



Another first from Amersham

L-α-Phosphatidylcholine, β-[1-¹⁴C]arachidonyl-
α'-stearoyl

CFA.655

- specifically labelled in β-position
- composition consistent with naturally occurring phosphatidylcholines
- >50mCi/mmol

Applications include

- membrane biochemistry
- studies of prostaglandin synthesis and function

Call Amersham for further details of our extensive range of phospholipids.

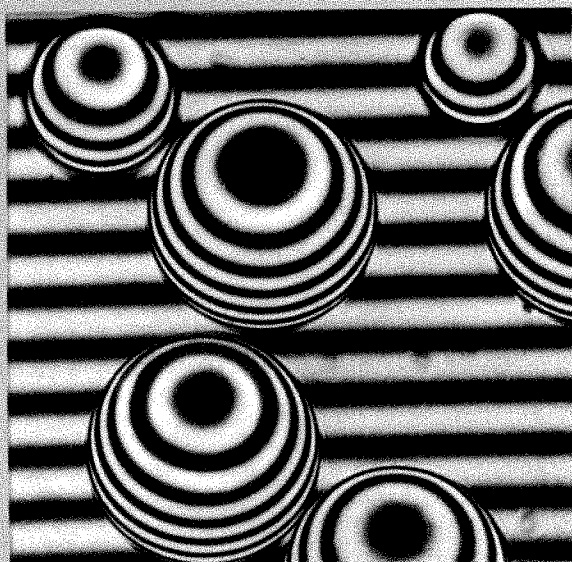
**The Radiochemical Centre
Amersham**

Circle No.13 on Reader Enquiry Card.

The Radiochemical Centre Limited, Amersham, England. Telephone: 024-04-4444.
In the U.S.A. and Canada: Amersham Corporation, Illinois 60005. Telephone: 312-364-7100
and 800/323-9750 (Tollfree).
In W. Germany: Amersham Buchler GmbH & Co KG, Braunschweig. Telephone: 05307-4691.

Go with Eupergit!

**The NEW carrier for immobilization
of biomolecules for biocatalysis
and affinity chromatography.**



TM = Trade Mark

- electroneutral, hydrophilic matrix with electroneutral, covalent binding mechanism
- high binding capacity due to high epoxide group content
- macroporous structure with a low water uptake
- high mechanical stability for large-scale industrial processes. Excellent flow properties and good filtration and centrifugation characteristics
- high chemical stability of the matrix and of the bonded ligands
- water uptake independent of pH and ionic strength (from pH 0-14)
- capability of reacting with low-molecular weight ligands in water or in organic solvents

Eupergit - your real choice!

Please send more detailed informations about
EUPERGIT to

name

street

city

country

Röhm Pharma
GMBH WEITERSTADT

Postfach 4347
D-6100 Darmstadt 1
West-Germany

Circle No.18 on Reader Enquiry Card.

nature

Volume 290 No.5803 19 March 1981

US versus science education . . .	175	Swiss nuclear power: Sanctions bite	180
. . . and UK versus universities	175	More interferon cloned	180
Kiss of life for Clinch River breeder: Reagan ditches Carter policy on proliferation	177	Indian community health: Cure for all ills	180
British academics up in arms	177	Oil tanker hazards: Inaction at sea	181
German big science: Win some, lose some	178	US animal research: Stricter safeguards	181
Madrid conference: Detente denied	178	UK animal research: Legislation unlikely	182
Moscow seminar threat	179	Commonwealth meeting: Food for reserves	182
Science education: Tap turned off	179	Romanian agriculture: Research expansion	182
Argentinians face charges	179	Correspondence: Genesis of genetics (K.F. Lindahl); Ethics of genes (I. Nabi); Cancer causation (E.M. Whelan); Badger debate (Lord Zuckerman; F. Yates; P.H. Roberts)	183, 184

NEWS AND VIEWS

Propagating rifts in sea floor spreading patterns: Tanya Atwater explains Richard Hey's model	185	Chondritic meteorites: R. T. Dodd explains their origin	189
Cellular transforming genes: Peter W.J. Rigby on the detection of genes involved in carcinogenesis	186	Why is the moon slowing down?: David W. Hughes describes the causes of lunar recession	190
Proline and protein procrastination: Roger H. Pain looks at the control of protein folding	187	Intergenic conversion and reiterated genes: Richard Egel on a novel means for exchange in the genome	191
Defining a lectin: Jan Kocourek and Václav Hořejší suggest a new definition	188	Computers in the new laboratory — A Nature survey	193

REVIEW ARTICLE

R. Peto, R. Doll, J.D. Buckley and M.B. Sporn	Can dietary beta-carotene materially reduce human cancer rates?	201
--	--	------------

ARTICLES

G.F. Davies	Earth's neodymium budget and structure and evolution of the mantle	208	B.M. Wilkins, G.J. Boulnois and E. Lanka	A plasmid DNA primase active in discontinuous bacterial DNA replication	217
S.Fields, G. Winter and G.G. Brownlee	Structure of the neuraminidase gene in human influenza virus A/PR/8/34	213	B. Jaurin, T. Grundström, T. Edlund and S. Normark	The <i>E.coli</i> β-lactamase attenuator mediates growth rate-dependent regulation	221

LETTERS

J.T. Clarke, H.W. Moos, S.K. Atreya and A.L. Lane	IUE detection of bursts of $\text{HLY}\alpha$ emission from Saturn	226	J.S. Patton, M.W. Rigler, P.D. Boehm and D.L. Fiest	Ixtoc 1 oil spill: flaking of surface mousse in the Gulf of Mexico	235
C.A. Leitch and J.V. Smith	Mechanical aggregation of enstatite chondrites from an inhomogeneous debris cloud	228	H.L.K. Simola, M.A. Coard and P.E. O'Sullivan	Annual laminations in the sediments of Loe Pool, Cornwall	238
G.E. Thompson and G.C. Wood	Porous anodic film formation on aluminium	230	W. Davison	Supply of iron and manganese to an anoxic lake basin	241
R.J. Francey	Tasmanian tree rings belie suggested anthropogenic $^{13}\text{C}/^{12}\text{C}$ trends	232	G. Topping and I.M. Davies	Methylmercury production in the marine water column	243

Cover: Photograph taken from the R/V G. W. Pierce of the Ixtoc 1 well blowout in the Gulf of Mexico (September 1979). The flames are covering a patch of ocean surface ~ 100 m across. (J. E. Barbash, Energy Resources Co. Inc.). The fate of the oil spill is described on p.235. This issue includes a feature on laboratory computers (p.193) and product review (p.xxiii).

M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd.

Contents continued overleaf

INBRED STRAINS IN BIOMEDICAL RESEARCH

MICHAEL F. W. FESTING

Geneticist at the Medical Research Council's
Laboratory Animals Centre, Carshalton.

£25; 495pp; ISBN 0 333 23809 5

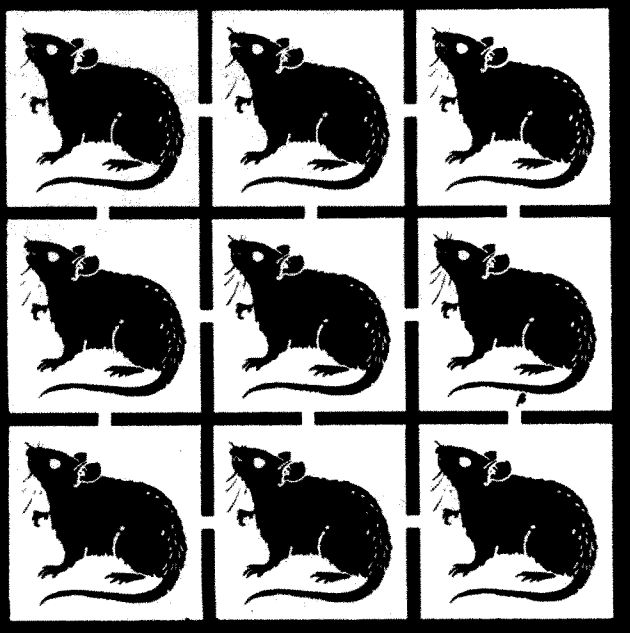
"This is an important work that should be on the shelves of every laboratory carrying out research with vertebrates." *Laboratory Animals*, January 1980.

Inbred strains of mice, rats and guinea pigs have been available for well over half a century and they have made a substantial contribution to many areas of biomedical research.

In spite of these major contributions, particularly in cancer research, there are still many areas of biomedical research in which inbred strains are hardly used at all, and even when they are used their properties are not fully understood. This is the first book to be devoted exclusively to inbred strains and their properties - over 400 are listed with details of their origin and known characteristics. This book will fill an important gap in laboratory animal literature and it is hoped that it will stimulate research workers in areas such as toxicology, biochemistry, parasitology and others, to consider ways in which inbred strains could improve the quality of their research.

M

For further details contact Frances Roach, Scientific and Medical Division,
Macmillan Publishers Ltd, Houndmills, Basingstoke, Hampshire.



Circle No. 30 on Reader Enquiry Card.

M

©1981 Macmillan Journals Ltd
Nature® is published weekly
ISSN 0028-0836

Registered as a newspaper at the
British Post Office

London

4 Little Essex Street, WC2R 3LF
Telephone: (01) 836 6633 Telex: 262024
Telegrams: Phusis London WC2R 3LF

Editor: John Maddox

Deputy Editor: Peter Newmark

Editorial Staff

Alun Anderson Sara Nash
Philip Campbell Peta Pickering
Isobel Collins Judy Redfearn
Konrad Guettler Miranda Robertson
Tim Lincoln Robert Walgate
Mary Lindley Charles Wenz
Naomi Molson Jonathan Wolfe

Publishing Director: Christopher Paterson

Marketing Director: Ray Barker

International Advertising Manager:

Andy Sutherland

Features Advertising Manager: Marion Delaney

New York

15 East 26 Street, New York, NY 10010
Telephone: (212) 689 5900

Robert Ubell (American Publisher)
Henry Dale (American Advertising Manager)
and Sheila Kane

Washington News Bureau

801 National Press Building, DC 20045
Telephone: (202) 737 2355 Telex: 64280
David Dickson (Washington News Editor)

Classified Advertising:

Jean Neville, 4 Little Essex Street, WC2R 3LF
Telephone: (01) 240 1101

Classified(USA)

Cathy Moore, 15 East 26 Street, New York, NY
10010 Telephone (212) 689 5900 Telex 668497

Display and Classified(Canada)

Peter Drake, 32 Front Street West,
201 Toronto, Ontario M5J 1C5
(416) 364 1623

Advertising Representatives in USA

Display

Nature New York Office—(212) 689 5900

Ad Craft (Houston)—(713) 627 9930

Didier & Broderick, Inc. (Chicago)—
(312) 498 4520

Jobson/Jordan/Harrison/Schulz, Inc.
(San Francisco) (415) 392 6794 (Los Angeles
& Pasadena) (213) 796 9200
Daniel Adams Associates, Inc.
Philadelphia—(215) 353 6191

CEL Associates (Boston)—(617) 848 9306
Brinker & Brinker (Fort Lauderdale)
(305) 771 0064

Japanese Advertising Representative

Orient Echo Inc.

11th Fl, Sankyo Bldg, 7-12, Tsukiji 2-Chome,
Chuo-Ku, Tokyo, Japan, Tel: 541-4923.

Annual Subscription including Index

UK & Eire	£75
USA & Canada	US\$173
Belgium	Airspeed only BF5695
West Germany	Airspeed only DM360
Netherlands	Airspeed only G390
Switzerland	Airspeed only SF330
Rest of Europe	Airspeed only £85
Rest of World	Surface £85
Rest of World	Airmail £130
(not USA, Canada & Europe)	

Back numbers (post-paid); UK, £2.00; USA &
Canada, US\$6.00 (surface), US\$8.50 (air); Rest
of World, £2.50 (surface), £3.50 (air).

Orders (with remittance) to:

Macmillan Journals Ltd, Brunel Road,
Basingstoke, Hants RG21 2XS.
Telephone: Basingstoke (0256) 29242
Telex: 858493

Vol. 290 No. 5802 12 March 1981

LETTERS

- | | | | | | |
|--|---|-----|---|---|-----|
| C. Gaven, C. Hillaire-Marcel
and N. Petit-Maire | A Pleistocene lacustrine episode
in southeastern Libya | 131 | P. Démant and D. Iványi | Further molecular complexities of
H-2 K- and D-region antigens | 146 |
| G.D. Floodgate, G.E. Fogg,
D.A. Jones, K. Lochte
and C.M. Turley | Microbiological and zooplankton
activity at a front in Liverpool
Bay | 133 | D.R. Hill and N.G. Bowery | ³H-baclofen and ³H-GABA bind
to bicuculline-insensitive GABA_B
sites in rat brain | 149 |
| P.A. Harper
and R.L. Juliano | Two distinct mechanisms of
fibroblast adhesion | 136 | T.N.R.V. Thampan
and J.H. Clark | An oestrogen receptor activator
protein in rat uterine cytosol | 152 |
| M.Cynader, F. Leporé
and J.-P. Guillemot | Inter-hemispheric competition
during postnatal development | 139 | G. Franchini, J. Even,
C.J. Sherr and
F. Wong-Staal | <i>onc</i> sequences (<i>v-fes</i>) of Snyder-
Theilen feline sarcoma virus are
derived from noncontiguous
regions of a cat cellular gene
(<i>c-fes</i>) | 154 |
| D. Beach and P. Nurse | High-frequency transformation
of the fission yeast
<i>Schizosaccharomyces pombe</i> | 140 | W.W. de Jong
and L. Rydén | Causes of more frequent deletions
than insertions in mutations and
protein evolution | 157 |
| H. Joenje, F. Arwert,
A.W. Eriksson, H. de Koning
and A.B. Oostra | Oxygen-dependence of
chromosomal aberrations in
Fanconi's anaemia | 142 | E. Fellenius, T. Berglindh,
G. Sachs, L. Olbe,
B. Elander, S.-E. Sjöstrand
and B. Wallmark | Substituted benzimidazoles inhibit
gastric acid secretion by blocking
(H⁺ + K⁺) ATPase | 159 |
| J.L. Grun and W.P. Weidanz | Immunity to <i>Plasmodium
chabaudi adami</i> in the
B-cell-deficient mouse | 143 | S.H. White, G.I. King
and J.E. Cain | Location of hexane in lipid
bilayers determined
by neutron diffraction | 161 |
| H.E. Blythman, P. Casellas,
O. Gros, P. Gros, F.K. Jansen,
F. Paolucci, B. Pau
and H. Vidal | Immunotoxins: hybrid molecules
of monoclonal antibodies and a
toxin subunit specifically kill
tumour cells | 145 | | | |

MATTERS ARISING

- | | | | | | |
|------------------------------|---|-----|---------------|--|-----|
| D. Rawlins; E.M. Standish Jr | Galileo's observation of
Neptune | 164 | E.A. Witalis | Plasmoid confinement by the
charged particle micro-fields | 166 |
| C.T. Kowal | Reply to Rawlins and Standish | 165 | G.C. Dijkhuis | Reply to Witalis | 166 |
| J.J.D. Greenwood | Does spider feeding behaviour
optimize dietary essential amino
acid composition? | 165 | | | |
| M.H. Greenstone | Reply to Greenwood | 165 | | | |

BOOK REVIEWS

- | | | | | | |
|-------------------|--|-----|------------------|---|-----|
| Stuart Sutherland | Brains, Machines and Persons
(by D. MacKay) | 167 | Derek Roe | Early Man in Britain and Ireland
(by A. Morrison) | 168 |
| Michael Bradbury | The Cerebral Microvasculature
(edited by H.M. Eisenberg
and R.L. Suddith) | 167 | Herbert Morawetz | Polymers (edited by R.A. Fava) | 169 |
| R.J.G. Savage | Pleistocene Mammals of North
America (by B. Kurtén and
E. Anderson) | 168 | Arthur W. Rose | Environmental Geochemistry:
A Holistic Approach
(by J.A.C. Fortescue) | 170 |
| | | | Otto Kandler | Microbiological Classification
and Identification (edited by
M. Goodfellow and R.G. Board) | 170 |

MISCELLANY

- | | | | |
|----------------|-----|------------------------|-----|
| 100 years ago | 90 | Announcements | 174 |
| Books received | 171 | Classified advertising | xi |

nature

Volume 290 No.5807 16 April 1981

Reagan must live with arms control	533	Safety at Windscale: Yes we were wrong	537
Can British education be reformed?	534	US university regulations: Complaints listed	537
European biotechnology lies in disarray: Commission's plan lacks friends	535	French nuclear power: Close to the brink?	538
German physics in row about Einstein	535	Nuclear Structure Facility: More delays ahead	538
Taking note, or shelving?	536	Correspondence: Genetic weaponry? (A. Lwoff); NRPB and risk (W. Blair, R.C. Thompson & H.M. Parker); Conservation sites (D.A. Radcliffe); Selling Darwin (G. Dover); From the museum . . . (P.H. Greenwood & J.F. Peake; C. Hill)	539, 540
European Space Agency: Space on the cheap	536		
Quistgaard awaits orders	536		
Cystic fibrosis: Diagnostic hopes	536		

NEWS AND VIEWS

Tissue specific gene expression: R.A. Flavell describes the control of the α amylase gene	541	Spinning asteroids: David W. Hughes describes ways of measuring asteroid shape	545
Rabies virus: J.S. Porterfield explains the role of antibody-mediated enhancement	542	Prolactin: Alan D.B. Malcolm looks at production and processing	546
Antibodies and introns: Miranda Robertson and Mike Hobart on versatility in the synthesis of immunoglobulins	543	Tropical soils: Ian Smalley on the problems of transferring soil technology to the developing world	547
Anomalous sputtering at high energy: P.K. Haff and L.E. Seiberling describe atomic scale erosion processes	544		

REVIEW ARTICLE

J.T.F. Zimmerman	Dynamics, diffusion and geomorphological significance of tidal residual eddies	549
------------------	---	-----

ARTICLES

G. Beamson, H.Q. Porter and D.W. Turner	Photoelectron spectromicroscopy	556	Y. Kurosawa, H. von Boehmer, W. Haas, H. Sakano, A. Trauneker and S. Tonegawa	Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes	565
H. Sakano, Y. Kurosawa, M. Weigert and S. Tonegawa	Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes	562			

LETTERS

R. Narayan and M. Vivekanand	A lower limit for the birth rate of pulsars	571	R.G. Carr and G.A. Somorjai	Hydrogen production from photolysis of steam adsorbed onto platinized SrTiO₃	576
F. Nagase, S. Hayakawa, H. Kunieda, F. Makino, K. Masai, Y. Tawara, H. Inoue, K. Koyama, K. Makishima, M. Matsuoka, T. Murakami, M. Oda, Y. Ogawara, T. Ohashi, N. Shibazaki, Y. Tanaka, I. Kondo, S. Miyamoto, H. Tsunemi and K. Yamashita	Spin-down episode of the X-ray pulsar Vela X-1	572	T.J. Shepherd and D.P.F. Darbyshire	Fluid inclusion Rb-Sr isochrons for dating mineral deposits	578
R.A. Stern, P.C. Agrawal and G.R. Riegler	The very soft X-ray transient H1853 + 60	573	S.C. Brassell and G. Eglinton	Biogeochemical significance of a novel sedimentary C₂₇ stanol	579
L.R. Fisher, R.A. Gamble and J. Middlehurst	The Kelvin equation and the capillary condensation of water	575	K. Pye	Rate of dune reddening in a humid tropical climate	582

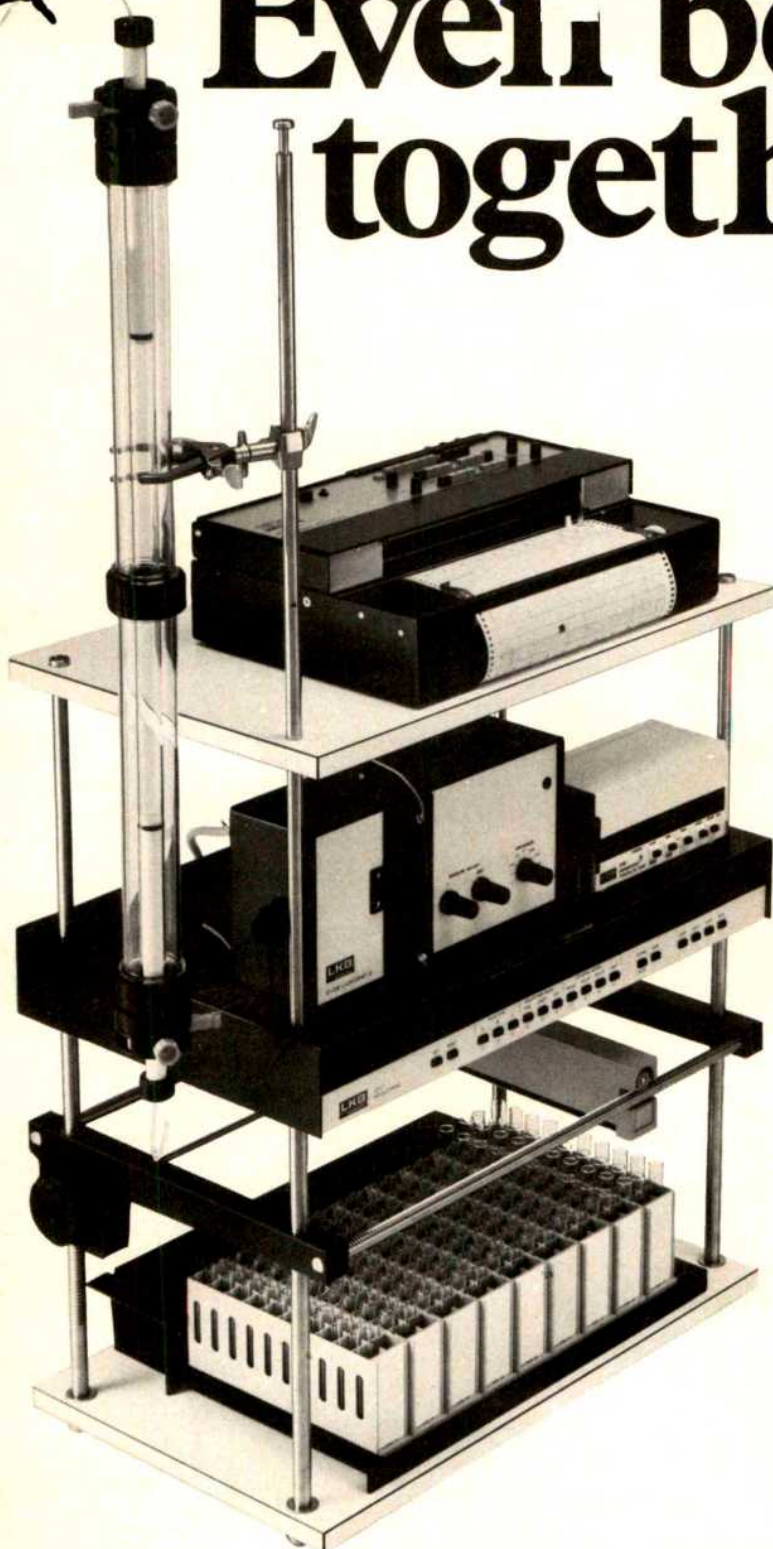
Contents continued overleaf

Cover: A thin section through a nerve fibre of a crayfish abdominal ganglion. Axonal and glial cell plasma membranes fuse with each other to form a well defined pore. See page 597.

M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to: *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd.

Even better together.



Individually, LKB chromatography instruments provide outstanding reliability and consistently accurate results.

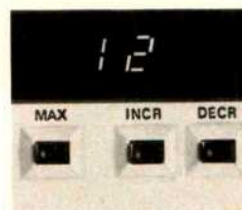
But together, they form total, integrated systems that are unequalled in their ease of use and dependability. The LKB 2111 MultiRac® Fraction Collector is the core of just such a system.

LKB excellence in design

The LKB 2111-MultiRac® system combines the ultra-reliable 2111 Fraction Collector with the new MicroPerpex® peristaltic pump, 2210 dual channel recorder and high sensitivity 2138 Uvicord®S UV monitor in one compact package. This entire system takes up less than 2 square feet of bench or cold room space. Yet, every instrument is easy to reach and adjust.

Integrated ease of operation

With the programmable 2111 MultiRac controlling the run, you're assured of accurate, no waste, no problem collections. And with precise pump control of fraction volumes,



automatic flow stop function during tube changes and continuous display of flow rate on the MicroPerpex pump, the system needs minimal interaction time with laboratory personnel.

Versatile and reliable

This new LKB collector draws on years of design experience and incorporates the most advanced circuitry. Because

the MultiRac precisely shifts the drop head, it can accommodate a large variety of collection vessels.

The sophisticated microprocessor controller of the collector lets you select time, drop or precise volume modes, diverts void volumes to waste and can be programmed to save tubes by collecting valleys in large volumes, peaks in small.



See it all together

Ask your LKB representative for more information on LKB complete chromatography systems. See how they work even better together.

LKB

Head office:
LKB Produkter AB
Box 305
S-161 26 Bromma, Sweden
Tel. 08/98 00 40, telex 10492

UK sales office:
LKB Instruments, Ltd.
232 Addington Road
South Croydon, Surrey CR28YD
England

Other sales offices in:
Athens (for the Middle East),
Copenhagen, Ghent,
The Hague, Hong Kong,
London, Munich, Paris,
Rome, Turku, Vienna

Circle No. 14 on Reader Enquiry Card.

J.E. Guest and A.M. Duncan	Internal plumbing of Mount Etna	584
W.J. Dominey	Anti-predator function of bluegill sunfish nesting colonies	586
K.W. Rammohan, H.F. McFarland and D.E. McFarlin	Induction of subacute murine measles encephalitis by monoclonal antibody to virus haemagglutinin	588
B.S. Prabhakar and N. Nathanson	Acute rabies death mediated by antibody	590
L.M. Nadler, P. Stashenko, R. Hardy, K.J. Tomaselli, E.J. Yunis, S.F. Schlossman and J.M. Pesando	Monoclonal antibody identifies a new Ia-like (p29,34) polymorphic system linked to the HLA-D/DR region	591
R.L. Cohen, K.A. Muirhead, J.E. Gill, A.S. Waggoner and P.K. Horan	A cyanine dye distinguishes between cycling and non-cycling fibroblasts	593
K. Kamino, A. Hirota and S. Fujii	Localization of pacemaking activity in early embryonic heart monitored using voltage-sensitive dye	595

C. Peracchia	Direct communication between axons and sheath glial cells in crayfish	597
O. Binah and Y. Palti	Potassium channels in the nodal membrane of rat myelinated fibres	598
J. Dorrington and R.E. Gore-Langton	Prolactin inhibits oestrogen synthesis in the ovary	600
L. Hall, R.K. Craig, M.S. Davies, D.N.L. Ralphs and P.N. Campbell	α-Lactalbumin is not a marker of human hormone-dependent breast cancer	602
B.C. Das and T. Sharma	Enhanced frequency of chromosome aberrations in lymphocytes of male compared with female muntjacs after X-ray irradiation <i>in vitro</i>	604
S.E. Kearsey and I.W. Craig	Altered ribosomal RNA genes in mitochondria from mammalian cells with chloramphenicol resistance	607

MATTERS ARISING

W. Leutenegger	Sexual dimorphism in early anthropoids	609	W.F. Eanes	Enzyme heterozygosity and morphological variance	609
J.G. Fleagle, R.F. Kay and E.L. Simons	Reply to Leutenegger	609	P. Handford	Reply to Eanes	610
			M.S. Starr and I.C. Kilpatrick	Nigral-derived muscimol circling—ipsiversive, contraversive or controversial?	610

BOOK REVIEWS

Ian Smart	Energy and Security (edited by D.A. Deese and J.S. Nye)	611	John Postgate	Methods for Evaluating Biological Nitrogen Fixation (edited by F.J. Bergersen)	613
W.B. Hubbard	Interiors of the Planets (by A.H. Cook)	612	Werner Haas	Regulatory T Lymphocytes (edited by B. Pernis and H.J. Vogel)	613
A. Rupert Hall	Galileo and the Art of Reasoning: Rhetorical Foundations of Logic and Scientific Method (by M.A. Finocchiaro)	612	Stuart Sutherland	The Psychobiology of Consciousness (edited by J.M. and R.J. Davidson)	614

MISCELLANY

100 years ago	548	Announcements	xv
Books received	615	Classified advertising	xvi

GUIDE TO AUTHORS

• Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.

• Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.

• Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.

• 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov. (1976).

^{32}P nucleotides you simply open and use



Our $> 600\text{Ci}/\text{mmol}$ dNTPs, $[\alpha - ^{32}\text{P}]$ - are now also available concentrated to $\sim 10\text{mCi}/\text{ml}$ in an aqueous solution, ready to pipet directly into your reaction mixture. No need to remove or concentrate the packaging solution. No unnecessary handling.

What's more, we've extended their radiochemical stability significantly with 10mM Tricine, a buffer known to be compatible in research systems. Biological testing by nick translation verifies their biological activity and assures you of reliable performance.

NEG-012A dATP, $[\alpha - ^{32}\text{P}]$ - NEG-014A dGTP, $[\alpha - ^{32}\text{P}]$ -
NEG-013A dCTP, $[\alpha - ^{32}\text{P}]$ - NEG-005A dTTP, $[\alpha - ^{32}\text{P}]$ -
 $> 600\text{Ci}/\text{mmol}$, $\sim 10\text{mCi}/\text{ml}$, in 10mM aqueous Tricine buffer solution (pH 7.6), Combi-V-Vial

Produced weekly and shipped daily. Send for complete ordering information on these and our standard concentration dNTPs, $[\alpha - ^{32}\text{P}]$ - .

Not for use in humans or clinical diagnosis

Circle No. 06 on Reader Enquiry Card.



New England Nuclear

549 Albany Street, Boston, Massachusetts 02118

Call toll free: 800-225-1572

(In Massachusetts and International: 617-482-9595)

NEN Chemicals GmbH: D-6072 Dreieich, W. Germany, Postfach 401240, Tel. (06103) 85034, Telex 4-17993 NEN D
NEN Canada: 2453 46th Avenue, Lachine, Que. H8T 3C9, Tel. 514-636-4971, Telex 05-821808

Tenth birthday for the environment	617	Nuclear reactor safety: Go for Super-Sara	621
Going private	617	Election in Israel: New deal for science?	621
Supersonic scandal	618	Kamikaze to Halley	621
Telecommunications monopoly eroded? Profiting from leased lines made seemly	619	Caribbean anti-pollution plan: Jamaican entente	622
Postgraduate medical schools in danger	619	Philosophy wrangle	622
Environmental asbestos: Dispute on levels	620	Correspondence: Irish evidence (W H G Rees), Who is Nabi? (E O Wilson), Darwin's truths (A Montagu, G Smith), Room for all (J R Baker); Popper's philosophy (A L Caplan), Academia stagnant? (J R Penswick), Freedom from NATO (R R C New), Conservation sites (K A Roberts)	623, 624
Spinning off research	620		
Space shuttle: Bonding problems	620		

NEWS AND VIEWS

Genes of lymphocytes: Miranda Robertson on the diverse means of generating antibody diversity	625	Strategies of viral oncogenesis: John Wyke	629
The expanding Earth: A D Stewart reports a recent conference	627	Measuring atmospheric gases and aerosols: Stuart A Penkett describes the GAMETAG experiment	631
Plesiosaurs — rigging and ballasting: Michael A Taylor explains why plesiosaurs swallowed stones	628	Cracks in rocks under stress: Barry Atkinson on the importance of microcracks	632

SPRING BOOKS SUPPLEMENT

Reviewers without pride or prejudice		633	G Y Craig	The Abyss of Time: Changing Conceptions of the Earth's Antiquity after the Sixteenth Century (by Claude C. Albritton Jr)	641
Harvey Brooks	Science at the White House: A Political Liability (by Edward J. Burger Jr)	634	William H McNeill	The Fates of Nations: A Biological Theory of History (by Paul Colinvaux)	642
J Z Young	Teaching and Learning about Science and Society (by John Ziman)	636	D T Donovan	Alfred Waterhouse and the Natural History Museum (by Mark Girouard)	644
Stuart Sutherland	Intelligence: The Battle for the Mind (H.J. Eysenck versus Leon Kamin)	636	A J Cain	The Evolutionary Synthesis: Perspectives on the Unification of Biology (edited by Ernst Mayr and Wilham B. Provine)	645
Peter Newmark	An Imagined World (by June Goodfield)	637	Alec Panchen	The Age of Birds (by Alan Feduccia)	646
Colin Burgess	The Beaker Folk: Copper Age Archaeology in Western Europe (by Richard J. Harrison)	638	Alison Jolly	The Natural History of the Gorilla (by A.F. Dixon)	647
A Rosalie David	An X-Ray Atlas of the Royal Mummies (edited by James E. Harris and Edward F. Wente). Mummies, Disease and Ancient Cultures (edited by Aidan and Eve Cockburn)	639	J T Enright	Intertidal Invertebrates of California (by Robert H. Morris <i>et al.</i>)	647
John Sutton	The Geology of Europe (by Derek V. Ager)	640	Jerry Donohue	The Double Helix: Text, Commentary, Reviews, Original Papers (edited by Gunther S. Stent)	648

Cover Dense cover of the colonial ascidian *Botryllus schlosseri* with arborescent *Botryllodes diegensis*. The ability to compete for space with already established animals influences the choice of settlement sites in several marine invertebrate larvae (see p 700)

M

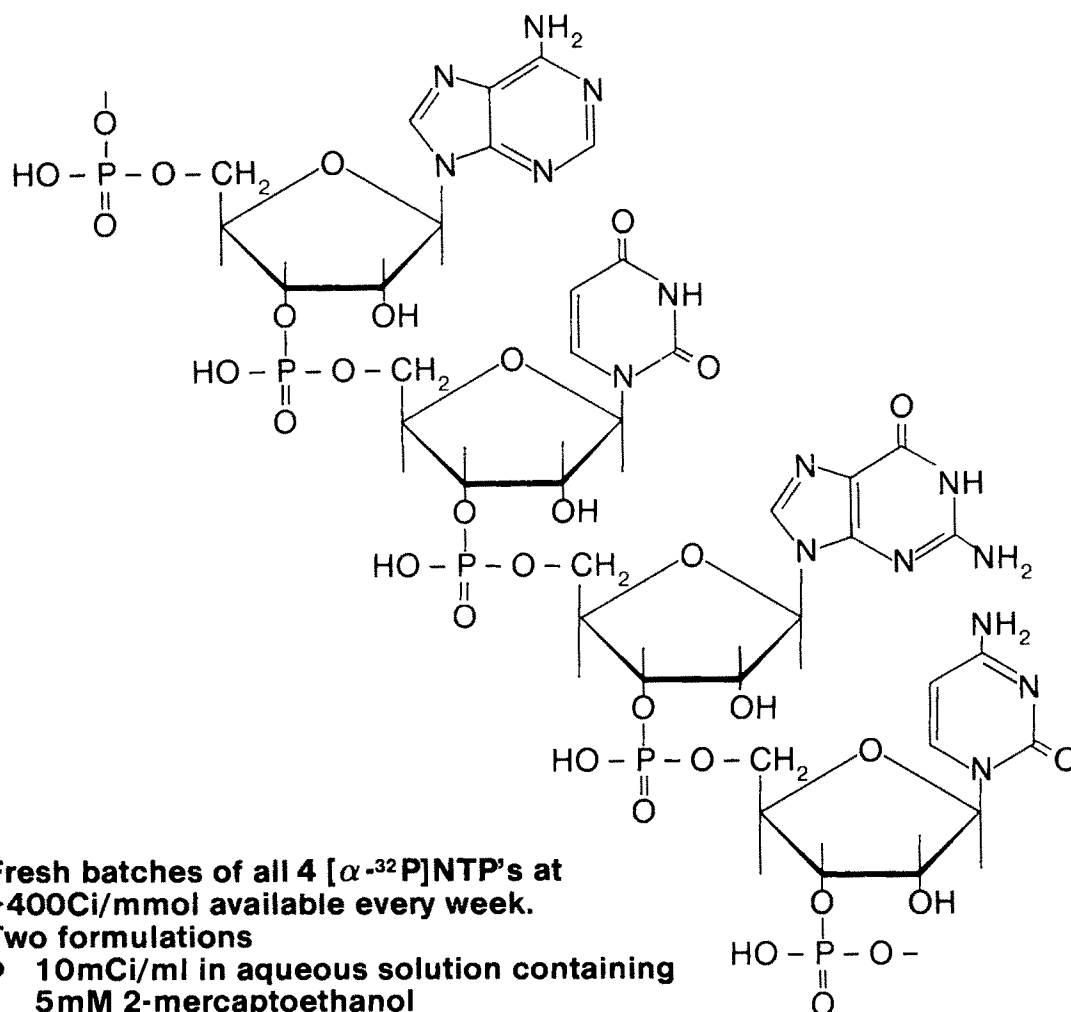
Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd

Contents continued overleaf

Ribonucleoside 5'-[α - ^{32} P] triphosphates

Straight from the bottle every week!

From 7th May 1981



Fresh batches of all 4 [α - ^{32} P]NTP's at >400Ci/mmol available every week.

Two formulations

- 10mCi/ml in aqueous solution containing 5mM 2-mercaptoethanol
- 1 mCi/ml in 50% aqueous solution

Radiochemical Purity

- All batches analysed by HPLC

Aqueous solutions at 10mCi/ml

- Supplied in "Duoseal V-vials"
- Reduced need for evaporation of the solvent
- Avoids the risk of decomposition during evaporation
- Reduced handling minimizes risk of wastage, spillage and contamination

Please ensure you have your copy of our despatch schedule

Call Amersham or your local representative for details.

**The Radiochemical Centre
Amersham**

The Radiochemical Centre Limited Amersham England Telephone 024 04 4444
In the USA and Canada Amersham Corporation Illinois 60005 Telephone 312 364 7100
and 800/323 9750 (Tollfree)
In W. Germany Amersham Buchler GmbH & Co KG Braunschweig Telephone 05307 4691

Eric Ashby	Small is Possible (by George McRobie)	650	Ann Henderson-Sellers	Rainbows, Halos, and Glories (by Robert Greenler)	655
James Gorman	Robert A. Heinlein: America as Science Fiction (by H Bruce Franklin)	651	John Stachel	The Greatest Power on Earth: The Story of Nuclear Fission (by Ronald W. Clark)	655
J Veverka	Voyage to Jupiter (by David Morrison and Jane Samz)	652	Paul Davies	Gravity, Black Holes, and the Universe (by Iain Nicolson)	656
David W Hughes	The Comet is Coming! The Feverish Legacy of Mr Halley (by Nigel Calder)	652	R H Dalitz	Encyclopedia of Physics (edited by Rita G. Lerner and George L. Trigg)	657
Brian O'Leary	Bound for the Stars (by Saul J. Adelman and Benjamin Adelman)	653	John C Marshall	Mindwatching (by H. and M. Eysenck)	657
Richard W Katz	The Rainmakers: American "Pluviculture" to World War II (by Clark C. Spence)	654	Clayton M White	The Audubon Society Encyclopedia of North American Birds (by John K. Terres)	658

REVIEW ARTICLE

J Mort and J Knights

Localization and electronic properties in amorphous semiconductors

659

ARTICLES

S F Dermott and C D Murray	Resonant structure of the asteroid belt	664	J Klysik, S M Sturdivant, J E Larson, P A Hart and R D Wells	Left-handed DNA in restriction fragments and a recombinant plasmid	672
D L Riddle, M M Swanson and P S Albert	Interacting genes in nematode dauer larva formation	668	A J Leigh Brown and D Ish-Horowicz	Evolution of the 87A and 87C heat-shock loci in <i>Drosophila</i>	677

LETTERS

F-L Deubner	Detection of low order p-modes in brightness fluctuations of the Sun	682	S-i Miyazaki and Y Igusa	Fertilization potential in golden hamster eggs consists of recurring hyperpolarizations	702
E Schatzman and A Maeder	Solar neutrinos and turbulent diffusion	683	T J C Jacob and G Duncan	Electrical coupling between fibre cells in amphibian and cephalopod lenses	704
G J Ashwell	Solvent inclusion and disorder effects on an isostructural series of organic conductors	686	T Deguchi	Rhodopsin-like photosensitivity of isolated chicken pineal gland	706
G Matsumoto, T Torii and T Hanya	High abundances of long-chain normal alkanolic acids in Antarctic soil	688	C K Omoto and G B Witman	Functionally significant central-pair rotation in a primitive eukaryotic flagellum	708
E I Hamilton	α -Particle radioactivity of hot particles from the Esk estuary	690	W Hsueh, U Desai, F Gonzalez-Crussi, R Lamb and A Chu	Two phospholipase pools for prostaglandin synthesis in macrophages	710
S C Brassell, A M K Wardroper, I D Thomson, J R Maxwell and G Eglinton	Specific acyclic isoprenoids as biological markers of methanogenic bacteria in marine sediments	693	W Gerhard, J Yewdell, M E Frankel and R Webster	Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies	713
M Schoell and F-W Wellmer	Anomalous ^{13}C depletion in early Precambrian graphites from Superior Province, Canada	696	C R Kahn, R Bertolotti, M Nino and M C Weiss	Short-lived cytoplasmic regulators of gene expression in cell hybrids	717
J Damuth	Population density and body size in mammals	699	M Katinka, M Vasseur, N Montreau, M Yaniv and D Blangy	Polyoma DNA sequences involved in control of viral gene expression in murine embryonal carcinoma cells	720
R K Grosberg	Competitive ability influences habitat choice in marine invertebrates	700			

Minerals and Rocks

Editors in Chief P J Wyllie

Editors A. El Goresy, W von Engelhardt, T Hahn

This series of monographs represents the continuation of 'Mineralogie und Petrographie in Einzeldarstellungen' on an international level. It presents up-to-date reviews and reports of carefully selected topics in mineralogy, geochemistry, petrology and isotope geology, including applications to industrial processes and products. Designed to supplement existing textbooks, many of these volumes can be used as the basis for advanced seminars.

A step towards the solution of the andesite problem

Volume 16 J. Gill

Orogenic Andesites and Plate Tectonics

1981 109 figures, 21 tables Approx 420 pages
Cloth DM 98,-, approx US \$ 46 70
ISBN 3-540-10666-9

The solution to the question of andesite genesis is a major, multidisciplinary undertaking facing geoscientists in the 1980's. **Orogenic Andesites and Plate Tectonics** was written in response to the growing need of researchers in this area for a clarification of the long-standing problem and identification of profitable areas for future investigations.

This book critically summarizes the vast relevant literature on the tectonics, geophysics, volcanology, geology, geochemistry, and mineralogy of andesites and their volcanoes. The author cites over 1100 references in these specialties and includes information on the location and rock composition of more than 300 recently active volcanoes. In addition, he provides a systematic and original evaluation of genetic hypotheses.

Numerous cross-references enhance the integrated subject development which consistently emphasizes the implications of data for theories of magma genesis. **Orogenic Andesites and Plate Tectonics** will prove an invaluable reference source to researchers and graduate students in the geosciences seeking a careful evaluation of genetic hypotheses and key data, arguments or gaps in this promising field.

Springer-Verlag
Berlin
Heidelberg
New York



Recent Volumes

Volume 15 J.B. Dawson

Kimberlites and Their Xenoliths

1980 84 figures, 35 tables XII, 252 pages
Cloth DM 88,-, approx US \$ 41 90
ISBN 3-540-10208-6

Volume 14 A.K. Gupta, K. Yagi

Petrology and Genesis of Leucite-Bearing Rocks

1980 99 figures, 43 tables XV, 252 pages
Cloth DM 69,50, approx US \$ 33 10
ISBN 3-540-09864-X

Volume 13 M.S. Paterson

Experimental Rock Deformation – The Brittle Field

1978 56 figures, 3 tables XII, 254 pages
Cloth DM 53,-, approx US \$ 25 30
ISBN 3-540-08835-0

Volume 9 J. Hoefs

Stable Isotope Geochemistry

2nd completely revised and up-dated edition
1980 52 figures, 23 tables XII, 208 pages
Cloth DM 59,-, approx US \$ 28 10
ISBN 3-540-09917-4

Prices are subject to change without notice

S977/4/1

nature

Volume 290 No 5809 30 April 1981

NEWS

New guard at the National Academy	723	Anti-viral drugs: Herpes on trial	727
Will the London Zoo survive (and how)?	724	European Earth sciences: Unity prevails	728
Vienna institute hit by espionage charges: Claims force Soviet official to resign	725	Out on a limb	728
Schuster quits as EEC research director	725	Interferon business: Trials ahead	728
US graduate employment: Industry has jobs	726	Dioxin hazards: Secrecy at Coalite	729
Soviet space programme: No shuttle yet?	726	Dioxin a sarcoma risk	729
Anglo-German project: Aurora held up	727	Correspondence: Majority verdict (J W Atz <i>et al</i>), 'I was there' (J L Gowans), Evolution's Waterloo (E O Wiley), No to sociobiology (B Banerjee)	730
US nuclear power: Cost disputed	727		

NEWS AND VIEWS

Springcleaning ribosomal genes: Gabriel Dover and Enrico Coen on a model for multigene evolution	731	Neutron optical activity: Leo Stodolsky on the first observations for matter waves	735
Continental margin processes: David G Roberts reports the Hedberg conference	733	Two DNA-binding proteins: David Davies describes the first structural studies	736
Solar physics at Taos: R Rosner describes results from the Solar Maximum Mission	734	Laminin: Brigid Hogan explains its role in epithelial cell attachment	737

ARTICLES

V M Canuto	The Earth's radius and the <i>G</i> variation	739	H Rojha, J R Miller, L C Woods and D M Glover	Arrangements and rearrangements of sequences flanking the two types of rDNA insertion in <i>D. melanogaster</i>	749
D B McKay and T A Steitz	Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA	744	W F Anderson, D H Ohlendorf, Y Takeda and B W Matthews	Structure of the <i>cro</i> repressor from bacteriophage λ and its interaction with DNA	754

LETTERS

R N Bracewell and W Graf	Solar quiescent prominences at 10.7 GHz	758	P J Lamb	Estimate of annual variation of Atlantic Ocean heat transport	766
B K Lucchitta, D M Anderson and H Shoji	Did ice streams carve martian outflow channels?	759	E Nakayama, H Tokoro, T Kuwamoto and T Fujinaga	Dissolved state of chromium in seawater	768
J-L Le Mouél, T R Madden, J Ducruix and V Courtillot	Decade fluctuations in geomagnetic westward drift and Earth rotation	763	J Oerlemans	Effect of irregular fluctuations in Antarctic precipitation on global sea level	770
D M Rowe, V S Shukla and N Savvides	Phonon scattering at grain boundaries in heavily doped fine-grained silicon-germanium alloys	765	J M Hurst and W S McKerrow	The Caledonian nappes of eastern North Greenland	772

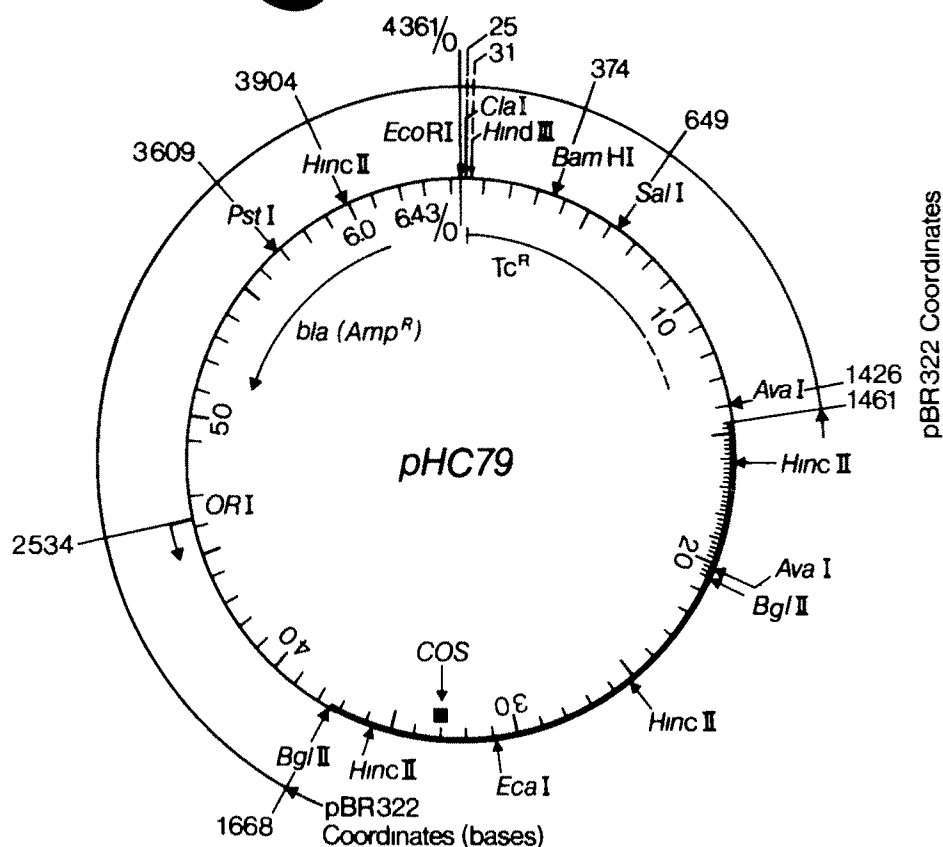
Contents continued overleaf

Cover: Flow channels around an 'island' *Kasei Valis*, Mars. On page 759, Lucchitta *et al* suggest that such channels are the result of ice flows, rather than liquid flooding

M

Nature® (ISSN 0028-0836) is published weekly, except the last week in December, by Macmillan Journals Ltd. Annual subscription for USA and Canada US \$173. Orders (with remittance) and change of address labels to: Macmillan Journals Ltd, Brunel Rd, Basingstoke RG21 2XS, UK. Second class postage paid at New York, NY 10010 and additional mailing offices. US Postmaster send form 3579 to *Nature*, 15 East 26 Street, New York, NY 10010. ©1981 Macmillan Journals Ltd

Cloning Vectors from Boehringer Mannheim



Hohn, B & Collins, J (1980) Gene 11, 291-298

	Cat No
Deoxyribonucleic acid, Cosmid pHC 79 from <i>Escherichia coli</i>	567 795
Deoxyribonucleic acid, Plasmid pBR 322 (Plasmid pBR 322 DNA) from <i>Escherichia coli</i>	481 238
Deoxyribonucleic acid, Plasmid Col E1 (Plasmid Col E1 - DNA)	208 370
Deoxyribonucleic acid, Plasmid Lambda dv1 (Plasmid Lambda dv1 - DNA)	208 345

and further 150 biochemicals for molecular biological research
are from stock available

please contact your local representative or



Boehringer Mannheim GmbH
Biochemica
P O Box 310120
D-6800 Mannheim
W Germany

J Cooke	Scale of body pattern adjusts to available cell number in amphibian embryos	775	R J Schneider, A Kulczycki Jr, S K Law and J P Atkinson	Isolation of a biologically active macrophage receptor for the third component of complement	789
G Morata and S Kerridge	Sequential functions of the bithorax complex of <i>Drosophila</i>	778	R Palacios and G Möller	Cyclosporin A blocks receptors for HLA-DR antigens on T cells	792
M C Fishman, P R Dragsten and I Spector	Immobilization of concanavalin A receptors during differentiation of neuroblastoma cells	781	J Rosen, T Ryder, H Ohtsubo and E Ohtsubo	Role of RNA transcripts in replication incompatibility and copy number control in antibiotic resistance plasmid derivatives	794
P Deterre, D Paupardin-Tritsch, J Bockaert and H M Gerschenfeld	Role of cyclic AMP in a serotonin-evoked slow inward current in snail neurones	783	O Huisman and R D'Ari	An inducible DNA replication-cell division coupling mechanism in <i>E.coli</i>	797
R M Sharpe, H M Fraser, I Cooper and F F G Rommerts	Sertoli-Leydig cell communication via an LHRH-like factor	785	J C Boothroyd, P E Highfield, G A M Cross, D J Rowlands, P A Lowe, F Brown and T J R Harris	Molecular cloning of foot and mouth disease virus genome and nucleotide sequences in the structural protein genes	800
F T Crews, S M Paul and F K Goodwin	Acceleration of β -receptor desensitization in combined administration of antidepressants and phenoxybenzamine	787			

MATTERS ARISING

D I Stirling and H Dalton	Fortuitous oxidations by methane-utilizing bacteria	803	E Volkin and P Benson	Cycloheximide inhibits DNA synthesis in SV40-transformed cells	804
I J Higgins, D J Best and R C Hammond	Reply to Stirling and Dalton	804			

BOOK REVIEWS

Marie Boas Hall	Never at Rest: A Biography of Isaac Newton (by R.S. Westfall)	805	Johannes Meienhofer	Polymer-supported Reactions in Organic Synthesis (by P. Hodge and D.C. Sherrington)	807
Douglas P Murray	American Science and Modern China 1876-1936 (by P. Buck)	806	John Friend	Plant Disease: An Advanced Treatise, Vol.5 (edited by J.G. Horsfall and E.B. Cowling)	808
Yves Farge	Synchrotron Radiation Research (by H. Winick and S. Doniach)	807	Robert Olby	Natural History Manuscript Resources in the British Isles (compiled by G.D.R. Bridson et al.)	808

MISCELLANY

100 years ago	738	Classified advertising	XII
Books	IX		

GUIDE TO AUTHORS

• Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted, it is advisable to consult us before writing a review article.

• Articles may be up to 3,000 words long with at most six displayed items (figures and tables), they are reports of major research developments.

• Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words, they have at most three or four displayed items.

• 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible. $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

- (i) only one reference number need be used if the reference is to several papers by identical authors
 - (ii) first and last pages of references should be cited
- Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated. 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material ideally, an unlettered original and three lettered copies should be provided, labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov (1976).

You need today's LANCET more than ever before. Here are a few reasons why.

1. THE LANCET is both timely and topical.

When an author submits an original article to THE LANCET, it is published with dispatch—usually in 4 to 8 weeks. After the issue is rushed to more than 50,000 subscribers around the world, comments and responses pour in from readers who speak their minds, knowing their views will get international attention. In 2 or 3 weeks their rebuttals, remarks, and reminders appear in the 30 to 40 columns of Letters to the Editor that are a regular feature of each LANCET. These letters keep you up to date on your colleagues' thoughts and actions, and give a distinctive personality to the journal.

THE LANCET

No. 4151		BOSTON, MASS. AND LONDON, SATURDAY 11th N 1988		Vol. 119, No. 1151	
ORIGINAL ARTICLES					
1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36
37	38	39	40	41	42
43	44	45	46	47	48
49	50	51	52	53	54
55	56	57	58	59	60
61	62	63	64	65	66
67	68	69	70	71	72
73	74	75	76	77	78
79	80	81	82	83	84
85	86	87	88	89	90
91	92	93	94	95	96
97	98	99	100	101	102
103	104	105	106	107	108
109	110	111	112	113	114
115	116	117	118	119	120
121	122	123	124	125	126
127	128	129	130	131	132
133	134	135	136	137	138
139	140	141	142	143	144
145	146	147	148	149	150
151	152	153	154	155	156
157	158	159	160	161	162
163	164	165	166	167	168
169	170	171	172	173	174
175	176	177	178	179	180
181	182	183	184	185	186
187	188	189	190	191	192
193	194	195	196	197	198
199	200	201	202	203	204
205	206	207	208	209	210
211	212	213	214	215	216
217	218	219	220	221	222
223	224	225	226	227	228
229	230	231	232	233	234
235	236	237	238	239	240
241	242	243	244	245	246
247	248	249	250	251	252
253	254	255	256	257	258
259	260	261	262	263	264
265	266	267	268	269	270
271	272	273	274	275	276
277	278	279	280	281	282
283	284	285	286	287	288
289	290	291	292	293	294
295	296	297	298	299	300
301	302	303	304	305	306
307	308	309	310	311	312
313	314	315	316	317	318
319	320	321	322	323	324
325	326	327	328	329	330
331	332	333	334	335	336
337	338	339	340	341	342
343	344	345	346	347	348
349	350	351	352	353	354
355	356	357	358	359	360
361	362	363	364	365	366
367	368	369	370	371	372
373	374	375	376	377	378
379	380	381	382	383	384
385	386	387	388	389	390
391	392	393	394	395	396
397	398	399	400	401	402
403	404	405	406	407	408
409	410	411	412	413	414
415	416	417	418	419	420
421	422	423	424	425	426
427	428	429	430	431	432
433	434	435	436	437	438
439	440	441	442	443	444
445	446	447	448	449	450
451	452	453	454	455	456
457	458	459	460	461	462
463	464	465	466	467	468
469	470	471	472	473	474
475	476	477	478	479	480
481	482	483	484	485	486
487	488	489	490	491	492
493	494	495	496	497	498
499	500	501	502	503	504
505	506	507	508	509	510
511	512	513	514	515	516
517	518	519	520	521	522
523	524	525	526	527	528
529	530	531	532	533	534
535	536	537	538	539	540
541	542	543	544	545	546
547	548	549	550	551	552
553	554	555	556	557	558
559	560	561	562	563	564
565	566	567	568	569	570
571	572	573	574	575	576
577	578	579	580	581	582
583	584	585	586	587	588
589	590	591	592	593	594
595	596	597	598	599	600
601	602	603	604	605	606
607	608	609	610	611	612
613	614	615	616	617	618
619	620	621	622	623	624
625	626	627	628	629	630
631	632	633	634	635	636
637	638	639	640	641	642
643	644	645	646	647	648
649	650	651	652	653	654
655	656	657	658	659	660
661	662	663	664	665	666
667	668	669	670	671	672
673	674	675	676	677	678
679	680	681	682	683	684
685	686	687	688	689	690
691	692	693	694	695	696
697	698	699	700	701	702
703	704	705	706	707	708
709	710	711	712	713	714
715	716	717	718	719	720
721	722	723	724	725	726
727	728	729	730	731	732
733	734	735	736	737	738
739	740	741	742	743	744
745	746	747	748	749	750
751	752	753	754	755	756
757	758	759	760	761	762
763	764	765	766	767	768
769	770	771	772	773	774
775	776	777	778	779	780
781	782	783	784	785	786
787	788	789	790	791	792
793	794	795	796	797	798
799	800	801	802	803	804
805	806	807	808	809	810
811	812	813	814	815	816
817	818	819	820	821	822
823	824	825	826	827	828
829	830	831	832	833	834
835	836	837	838	839	840
841	842	843	844	845	846
847	848	849	850	851	852
853	854	855	856	857	858
859	860	861	862	863	864
865	866	867	868	869	870
871	872	873	874	875	876
877	878	879	880	881	882
883	884	885	886	887	888
889	890	891	892	893	894
895	896	897	898	899	900
901	902	903	904	905	906
907	908	909	910	911	912
913	914	915	916	917	918
919	920	921	922	923	924
925	926	927	928	929	930
931	932	933	934	935	936
937	938	939	940	941	942
943	944	945	946	947	948
949	950	951	952	953	954
955	956	957	958	959	960
961	962	963	964	965	966
967	968	969	970	971	972
973	974	975	976	977	978
979	980	981	982	983	984
985	986	987	988	989	990
991	992	993	994	995	996
997	998	999	1000	1001	1002
1003	1004	1005	1006	1007	1008
1009	1010	1011	1012	1013	1014
1015	1016	1017	1018	1019	1020
1021	1022	1023	1024	1025	1026
1027	1028	1029	1030	1031	1032
1033	1034	1035	1036	1037	1038
1039	1040	1041	1042	1043	1044
1045	1046	1047	1048	1049	1050
1051	1052	1053	1054	1055	1056
1057	1058	1059	1060	1061	1062
1063	1064	1065	1066	1067	1068
1069	1070	1071	1072	1073	1074
1075	1076	1077	1078	1079	1080
1081	1082	1083	1084	1085	1086
1087	1088	1089	1090	1091	1092
1093	1094	1095	1096	1097	1098
1099	1100	1101	1102	1103	1104
1105	1106	1107	1108	1109	1110
1111	1112	1113	1114	1115	1116
1117	1118	1119	1120	1121	1122
1123	1124	1125	1126	1127	1128
1129	1130	1131	1132	1133	1134
1135	1136	1137	1138	1139	1140
1141	1142	1143	1144	1145	1146
1147	1148	1149	1150	1151	1152
1153	1154	1155	1156	1157	1158
1159	1160	1161	1162	1163	1164
1165	1166	1167	1168	1169	1170
1171	1172	1173	1174	1175	1176
1177	1178	1179	1180	1181	1182
1183	1184	1185	1186	1187	1188
1189	1190	1191	1192	1193	1194
1195	1196	1197	1198	1199	1200
1201	1202	1203	1204	1205	1206
1207	1208	1209	1210	1211	1212
1213	1214	1215	1216	1217	1218
1219	1220	1221	1222	1223	1224
1225	1226	1227	1228	1229	1230
1231	1232	1233	1234	1235	1236
1237	1238	1239	1240	1241	1242
1243	1244	1245	1246	1247	1248
1249	1250	1251	1252	1253	1254
1255	1256	1257	1258	1259	1260
1261	1262	1263	1264	1265	1266
1267	1268	1269	1270	1271	1272
1273	1274	1275	1276	1277	1278
1279	1280	1281	1282	1283	1284
1285	1286	1287	1288	1289	1290
1291	1292	1293	1294	1295	1296
1297	1298	1299	1300	1301	1302
1303	1304	1305	1306	1307	1308
1309	1310	1311	1312	1313	1314
1315	1316	1317	1318	1319	1320
1321	1322	1323	1324	1325	1326
1327	1328	1329	1330	1331	1332
1333	1334	1335	1336	1337	1338
1339	1340	1341	1342	1343	1344
1345	1346	1347	1348	1349	1350
1351	1352	1353	1354	1355	1356
1357	1358	1359	1360	1361	1362
1363	1364	1365	1366	1367	1368
1369	1370	1371	1372	1373	1374
1375	1376	1377	1378	1379	1380
1381	1382	1383	1384	1385	1386

5 March 1981

Bulley
25/2/82

Pope John Paul II in Hiroshima

The Pope's visit last week to non-Christian Japan showed how well this indefatigable traveller can tune his speeches to his audience. With the faithful thin on the ground, general issues took precedence over those that domestically occupy his Church. The result was a more open discussion of some perplexing problems than in the Philippines the week before. And nobody should complain that the Pope followed other statesmen to Hiroshima for a major speech on the problem of nuclear war. The melodrama of the site ensures that some notice will be taken of what is said. Not surprisingly, the Pope's discussion was a good deal more perceptive than many that have preceded it. The visit also provided him with an opportunity to pat on the back one of the two sponsors of the occasion, the United Nations University. The university's new rector, the Indonesian Soedjatmoko, seems well placed to help the university make a mark in its chosen field of interest, the development of developing countries.

The danger facing those who make speeches at Hiroshima is that of following the same hackneyed theme. Nuclear weapons are awful things, and the product of scientific research, from which it follows that the sharpest dilemma is the Promethean dilemma. The Pope fortunately had the wit to acknowledge that this task has political as well as ethical or moral dimensions. However powerful may be people's collective wish that nuclear war should not break out, their political differences ensure that the danger will remain alive. It would be asking too much to accept the argument, for which there is some justification, that only the balance of terror has kept the peace these past thirty years. But the implication of his remarks (at Nagasaki) that the United Nations must more powerfully assume a peacekeeping role is needlessly unearthly, the United Nations is a place where governments with different interests meet, as often to declare their differences as to resolve them. Although the United Nations has from time to time passed high-sounding resolutions on the need for "general and complete disarmament" (known as "GCD" in the trade) the painfully slow processes of the Committee on Disarmament shows what problems lie ahead.

Fortunately, not all problems are such inherently gloomy problems. Although the theme of the Pope's speech last week was "the ethical problems of the technological society", he reminded those who heard him that "we all greatly benefit" from the application of science and technology. In the past few gloomy years, optimism on this score seems to have melted surreptitiously away, not so much from anxiety about war as because of the faltering of the industrial economies and the predictable but nonetheless bitter failure of efforts such as the United Nations conference on technology and development two years ago. But can it be that the promise of science and technology for the improvement of the human condition, widely believed only a few years ago, has evaporated overnight? Is it not more likely that people have become discouraged? And, in these circumstances, may it not be especially valuable that vaguely incongruous figures like the Pope should go about reminding people that the promise is still there to be captured?

But how? At Hiroshima, the Pope said that the need was for a "moral about-face". Translated into the language of everyday politics, the implication is that governments must become more serious in their consideration of documents such as the report of the Brandt Commission on the problems of contemporary development, more imaginative in their management of domestic wealth and liberty and more courageous in their handling of dealings with each other. But, in the last resort, governments —

even illiberal governments — are merely the products of their people. Is the practical message of what the Pope was saying last week about the moral responsibility of the scientific community that the time has come when governments should be reminded again, as they have been *ad nauseam* in earlier decades since the Second World War, that the time has come to start solving problems, not to be mesmerized by them?

Ructions in space

The European Space Agency is using unaccustomed undiplomatic language in its attempts to persuade the new United States Administration not to force the cancellation of the joint venture to put a pair of spacecraft into polar orbits about the Sun. Rumours that this might be one of the consequences of the cut in the budget of the National Aeronautics and Space Administration have been circulating in Washington for several days (see *Nature* 26 February, p 738). The European reaction to the news that the rumours are true has been understandably tetchy. Europe, the European Space Agency is saying, has been led up the garden path, has been induced to lavish a substantial part of the meagre funds available for planetary exploration on a project whose value will now be much reduced (see page 3), and which may even have to be abandoned altogether. By asking that member states should send their diplomats into action in Washington, the European Space Agency has broken new ground in the foreign relations of multilateral agencies like itself. The bad temper in its public statement on the problem suggests that good relations may be at an end.

Luckily, that time may not yet have arrived. There remains a chance that the American half of the solar polar mission will be restored in the budget to be published on 10 March. More to the point, there is every chance that the United States Congress, still not quite certain what to make of President Reagan and his well-intentioned but draconian fiscal policies, will write back into the budget some of the items the President is planning to strike out. (And then, of course, it is always possible that the White House might decide not to spend the money, anyway.) In short, it is too soon to be supposing that the solar polar experiment is a washout. What has happened is more subtle but also more important.

Domestically, the United States space programme is in a mess. For the past three years, the space shuttle, the machine intended as the launching vehicle both for commercial and scientific spacecraft, has been in trouble. It is already two years late, and further difficulties may come to light in the critical months ahead. Yet it is too late to complain that the United States should have embarked on a quite different programme of development ten years ago, when the Apollo programme ended. Willynilly, the shuttle is the only sufficiently versatile launching system in sight. If its use is further delayed, the consequences will be more serious than the cancellation of the American half of the solar polar project. The European half may be without a launcher as well, while a whole string of communications satellites will have to find their way into synchronous orbits by quite different means. Worse still, for the scientific community, the Large Space Telescope in which the European Space Agency also has a substantial stake and which the United States Administration has so far (to its credit) kept free from cuts, will also be in jeopardy. Bad temper is, in the circumstances, inappropriate.

Looking further ahead, however, the European Space Agency

is right to be alarmed. In the past few years, it has dawned on European governments that they must have access to some way of launching spacecraft of all kinds. The chance to join the shuttle project as a partner has however gone. The European launcher Ariane may have a future, if only as an insurance policy against further delay of the shuttle and as an assurance of independence, but is not the ideal way of meeting the need. The space agency itself, now under new management, has a poor public reputation stemming at least in part from its bureaucratic ways and from the way in which member governments appear to regard it as their first objective to ensure that their contributions are required by contracts placed with national companies. Many scientists concerned to put scientific experiments into space have found it easier to work with the American agency than with that in Paris. The first lesson to be learned in Paris from the threatened cancellation of the American half of the solar polar project is that it should put its own house in order. Only then will it be able to take the lead in making sure that European governments collectively have a view of their future role in this important field.

The European Space Agency would at the same time do well to draw a wider lesson from what is going on in Washington. The space budget is being cut not only because of the technical problems with the shuttle but also because expenditure of this kind is not immediately productive. In the long run, it is better for Europe that the American economy should be healthy (free from inflation) and productive than that funds should be found for the second half of the solar polar project. And, it would be folly if the growing European enthusiasm for space projects were further to undermine the economic health of Western Europe.

Questions answered

The British predilection for studying the machinery for administering publicly supported science has burst out again, this time in the House of Lords (*Nature* 26 February, p 741). British governments seem constitutionally unable to decide for themselves what sensible steps to take, and usually set up *ad hoc* committees to tell them what to do. The select committee of the House of Lords is an unexpected but a better way of tackling the problem. The committee has also taken the unusual step of inviting written evidence from anybody, and the specific questions it has asked show that its inquiry may well be radical. What follows may help to start the ball rolling.

The most important, because the most novel, group of questions is that labelled machinery of science. Could the relationship between the British government and the scientific community be improved and, if so, how? The obvious difficulty is that, formally, there is no such thing as a scientific community. There are a few organizations, the Royal Society chief among them, which are sufficiently diverse to be consulted as if they were representative, and so eminent that the advice they give is rarely foolish. However radically the House of Lords committee tackles its inquiry, these influences on government will (and should) continue. Moreover, it would be fruitless to engage an army of constitutional lawyers to devise some way in which the wider scientific community could speak with even greater representative authority. The truth is that opinions within the scientific community differ about all the questions the committee has asked as well as about more specific problems.

What the British government needs is some means by which it can sense the diversity of scientific opinion on contentious issues, forming in the process a judgement about which views should be taken seriously, and at the same time making more effective use of the imagination and willingness to help of people working in laboratories of all kinds. The mechanism required is more like a lightning-rod than a forum. Whether it should be more like the United States President's Science Advisory Committee than the British Advisory Council on Scientific Policy of the 1950s is less important than that it should be accessible to all who want to tell it something.

Given such a device for drawing attention to important problems, some of the select committee's questions would be

answered. Thus, while no amount of organization can ensure that there is sufficient communication between those responsible for the administration of science in industry, research councils and government, common experience shows that public servants respond diligently to well-considered public complaint or even to constructive suggestion. To be effective over the whole field the select committee has mapped out, a committee of this kind would have to be adequately serviced — previous essays in this direction have been less than successful because they have been run on a shoestring. The terms of reference of the existing Advisory Council on Applied Research and Development are too narrow to do what is now required, and the council seems shy of tackling questions of how well government machinery is working. It is, however, essential that any such organization should be independent of particular government departments. The Council for Scientific Policy of the 1960s, which did some useful work, was too dependent on the Department of Education and Science to be effective. The need now is for a channel for communication free from departmental interests. The House of Lords is well placed to decide how such a committee should be plugged into the government machine. Direct responsibility to the Prime Minister would be the constitutional ideal but would make a mountain out of a molehill. A connection with the Treasury would not offend the lawyers and could be fun. Otherwise, the best choice would be a return to the 1960s, when the Lord President of the Council was responsible for the Council on Scientific Policy.

Other questions on the list from the House of Lords bear on the issue of the Rothschild customer-contractor principle. In spite of the mixed experience of the past decade, it would be wrong now to give up the notion that government departments should be equipped to decide scientific questions relevant to their own work, and to commission their own research. But the system in which departments were required to appoint chief scientists to manage their scientific affairs has not always worked well. Some of those appointed have not been up to the job, few of them have been in their posts long enough to learn what to do and none of them has been paid enough. However, there can be no general rule for deciding how much money should be spent by departments and how much by research councils. Research councils should have enough money to discharge their responsibilities properly, and government departments should spend money on research in the expectation of winning tangible benefit as a result. Some of the imperfections of the present system, and especially the difficulties which have arisen in the financing of applied research in which no government has a commanding interest, would not exist if the arrangements for dividing the science budget between those dependent on it took more account of the scientific merits of their different programmes.

These are the easy questions. The committee's question about the adequacy of the government's own scientific manpower is much more difficult. What should be done about the scientific civil service? And what should be done to equip the British civil service as a whole to grapple with modern problems? The British government is at present too directly involved in research in its own establishments (shockingly so in defence research). If the committee does more than skate round the manpower problem, it will recommend that the government reduce its own research effort drastically (but provide better careers for those who remain) and also tackle seriously the recruitment of scientists in mid-career to the civil service proper. The present system merely perpetuates amateurism and ensures inexpert advice.

The hardest question of all, however, is missing from the list circulated by the House of Lords. What is to be done about defence research? Only two weeks ago (*Nature* 19 February), the European Commission was pointing out that defence research is a larger proportion of all research in the United Kingdom than elsewhere in the European Community. Outsiders have no means of telling how necessary this vast enterprise may be, although there is every reason to believe that the huge defence laboratories could be used more effectively to foster industrial change without hazarding national defence. The select committee should think of adding this topic to its agenda.

Open row about joint space project

European space agency plans protest

Washington

Western European nations have agreed to organize a top-level protest to the US State Department at the proposal of the National Aeronautics and Space Administration (NASA) to withdraw its spacecraft from the International Solar Polar Mission.

The mission, under which two spacecraft would be flown simultaneously over opposite poles of the Sun, is being organized jointly by NASA and the European Space Agency. Originally each agency was to have supplied one of the spacecraft, the two being launched from the space shuttle in 1985. NASA has now decided to absorb some of the massive cuts in its space science programme required by the Reagan Administration by terminating the work on its own spacecraft. It would still provide support facilities for the European spacecraft if the European Space Agency decided to press on alone.

The European agency issued a strongly-worded statement last week saying that it considered NASA's actions to be a unilateral infringement of the memorandum of understanding signed between the two countries, and warning that it could have severe consequences for future co-operation in space research between Europe and the United States. Now several countries, including France, West Germany, Italy and the United Kingdom, have agreed to make a joint protest to the State Department, hoping that NASA can be persuaded to change its mind. Without the participation of the US spacecraft, which would have provided the first opportunity to study a star at close hand, the value of the mission will be severely limited.

Within NASA, however, officials have little hope that the proposed cuts will be reversed, since next to the space shuttle programme NASA's top priority has been to keep intact its two main space science projects, the space telescope and the Galileo mission to Jupiter.

NASA officials have stressed the importance they attach to maintaining links with European colleagues, as both acting deputy director Anthony Calio and the director of the Goddard Space Flight Center, Dr A. Thomas Young, pointed out when Mrs Margaret Thatcher visited the centre during her visit to Washington last week.

However, NASA has had to take some hard decisions itself, and apparently decided that the cancellation of its share in the solar mission was the least of several evils. On the other hand, the deferral of the

gamma-ray observatory and of the Venus Orbiting Imaging Radar have been demanded by the White House.

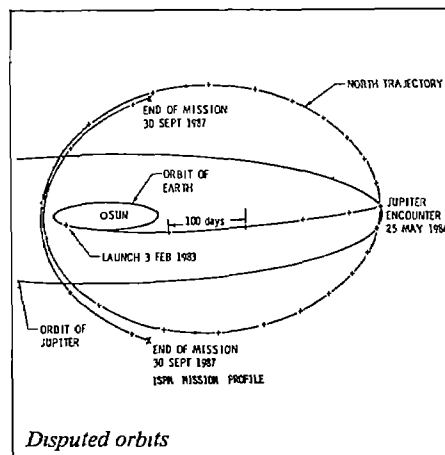
NASA has now been asked by the Office of Management and Budget (OMB) to make even more cuts in its next year's budget, following the discovery that the size of the required federal budget cut had been underestimated. Although it is not known where these most recent cuts will fall, one rumour is that funding for astrophysics experiments on the second Spacelab will be withdrawn, a decision which would particularly hit British scientists.

When rumours began to circulate in Washington last month that OMB was proposing to scrap Galileo, West Germany — which is developing several of the instruments to be flown on both the orbiter and the probe and is contributing about \$50 million to the cost of the mission — protested strongly to Secretary of State Alexander Haig. Mr Haig's subsequent intervention with OMB is said to have contributed significantly to the decision not to terminate the project.

No mention was made at that time of a possible threat to the solar polar mission. NASA was given its total required budget cut but left to decide where the cuts would fall. Cancelling the solar polar mission spacecraft would save between \$250 and \$300 million total costs over the next eight or nine years.

Gloom at prospect of cancellation

Officials at the European Space Agency (ESA) are too upset by the decision by the National Aeronautics and Space Administration (NASA) to ditch its solar-polar spacecraft to talk about contingency plans, preferring to await the results of diplomatic efforts to get the decision reversed before 10 March when the federal budget goes to Congress and after which changes cannot be made. Nevertheless it is almost certain that, if it has to, ESA will go it alone, having already invested too much in the mission to abandon it.



The European Space Agency council meets this Thursday (5 March) to decide on its reaction to the US moves. There is still a chance that some of the US instrumentation could be transferred to the European craft, although European experiments on the US craft would be given higher priority. The agency could equally well decide that, without full US participation, the mission is not worth pursuing, a decision which would severely disrupt its space science budget.

NASA's final proposals are expected to be announced next week, when Mr Reagan gives Congress a detailed list of his proposed budget reductions. There will undoubtedly be some opposition in Congress if the mission remains among the proposed cuts. Senator Jack Schmitt, an ex-astronaut and now chairman of the science and space subcommittee of the Senate Commerce Committee, said last week that he was very concerned that the decision could jeopardize future cooperation, particularly at a time when Europe seemed to be moving ahead of the United States in some areas of space technology.

But, as a loyal supporter of Mr Reagan, Mr Schmitt said he also acknowledged that there was a need to make substantial cuts in federal spending to restimulate the economy — and that space science, together with other research areas, would have to accept its fair share. **David Dickson**

The International Solar Polar Mission was originally conceived as a two-spacecraft mission that would observe dynamic and variable phenomena at high solar latitudes. The spacecraft's journey round Jupiter would also provide the first opportunity for spatially and temporally resolved measurements of Jupiter's magnetosphere. Some of the experiments to be flown on the two spacecraft would be identical but others would complement each other.

A decision to abandon the NASA spacecraft would mean that spatial and temporal resolution could be achieved less adequately only by comparison with simultaneous observations from the Earth. In all probability, it would also mean that the NASA spacecraft's ability to make optical measurements of the Sun's surface — unlike the ESA spacecraft, it has a pointing device — would be lost.

Each spacecraft is to house some of the other partner's experiments. It is thought unlikely, because of commitments already made, that the experiments to be flown on the ESA spacecraft could be changed at this late stage. Fortunately NASA has said that it will continue support of its experiments to be flown on ESA's craft. That, however, begs questions of the

future of European experiments for the NASA vehicle. According to ESA, work on those experiments is already well advanced and substantial sums of money, which will have been spent in vain if NASA's decision stands, have already been committed. The total cost to completion of the two spacecraft is an estimated \$140 million, \$100 million of which is already accounted for.

Judy Redfearn

European fusion

JET's ambition

JET, the joint European tokamak nuclear fusion experiment, may cost another 400 million European units of account (EUA) (£216 million) if first the JET Council — which oversees the project — and then the European Commission agree. The money would be spent over five years, beginning in 1982, and would extend JET to its full design performance, including bulk ignition of a deuterium-tritium plasma. "If they want the results, they'll have to pay" said Dr Hans-Otto Wüster, director of JET, of the European governments last week.

The original cost of basic JET was 185 million EUA at January 1977 prices. Inflation in Britain, coupled with the strength of the pound sterling and a serious underestimate of the cost of diagnostic equipment, has pushed that basic price to something nearer 300 million EUA, but even so the proposed extension would more than double the total cost. Dr Wüster, however, points out that the extra cost is only twice the annual budget of the CERN laboratory in Geneva.

The rush to extend JET is occasioned by two factors. First, machines like JET seem to have a better chance of reaching ignition, the nuclear burning of the plasma, than when they were designed in the early 1970s. Second, the European Council of Ministers has asked for a proposal for fusion research in Europe for the next five-year period, 1982–86, by 1 July this year.

Competition with the United States is also in mind. Princeton's Tokamak Fusion Test Reactor is due to begin experiments in 1982, and the somewhat larger JET in early

1983. Princeton is being pushed hard by the US Department of Energy to try ignition early. If all went well, JET's initial non-nuclear plasma experiments would then look pretty tame. On the other hand, if either laboratory is too hasty in introducing radioactive tritium, repeating early experiments will be very difficult, requiring remote handling and radiation protection.

Important plasma physics problems remain to be solved. Present optimism arises because two predicted limits on plasma containment have not been observed in existing small tokamaks. "Trapped particle instabilities" were expected to set in at temperatures above 30 million K, but the smaller Princeton tokamak has reached 70 million K without seeing them. And the ratio of plasma pressure to magnetic pressure was thought to be limited to about 1 per cent, but values of 8–9 per cent have been reached. While experimentally encouraging, these results are theoretically puzzling, so the JET and Tokamak Fusion Test Reactor regions may still hold surprises.

Dr Wüster is therefore cautious, offering 3–4 years of plasma at JET before attempting ignition around 1987. But the bulk of the spending on "extended performance" would have to be undertaken in 1982–86. In January, the JET Council gave Wüster a 400 million EUA guideline for the extended performance, and asked him to say what he could do with that. If, this month, his proposals are adopted, it will be for the European Commission to consider whether to include it as part of the 1982–86 research programme.

Conflicts between the interests of the national laboratories and money-hungry JET seem likely to be resolved at JET Council level, but the commission must also test the interest of governments, for which reason it has set up a new committee — the Consultative Committee for Fusion Programmes — which will be dominated by the representatives of government departments from whose coffers the money must ultimately be found.

The commission has also established an expert panel, the European Fusion Review Panel, to make recommendations for a long-term programme, beyond the present generation of machines. This panel will also report to the commission before July, in time to influence the final commission proposal.

One question to be dealt with is whether Europe is wise to commit itself so heavily to one design of fusion device, the tokamak. In the United States, there is also substantial work on the potentially simpler magnetic mirror confinement and inertial confinement. Wüster says that it makes sense to push ahead with the device that will most quickly yield ignition, so as to win practical experience of the problems of operating fusion reactors of all types. We shall see whether the European Fusion Review Panel agrees with him.

Robert Walgate

United States budget

More cuts

Washington

True to its political convictions, when full details of its proposed budget cuts are announced next week, the Reagan Administration is expected to eliminate virtually all the federally sponsored programmes initiated by President Carter to stimulate innovation in private industry.

President Carter's initiatives were the result of a broad-ranging, eighteen-month study largely instigated by the Office of Science and Technology Policy. Recommendations were made by eight independent advisory committees, and submitted to the White House with proposals from individual federal agencies.

From these, a package of 32 proposals was eventually accepted and announced by Mr Carter in November 1979. Although criticized at the time for not going far enough, it was generally agreed that the proposals made up a modest set of experimental and exploratory approaches aimed at bringing industry, government and the universities closer together.

The interventionist approach which the new programmes embodied, however, has found little sympathy with the new Administration, which sees its principal strategy for stimulating innovation as improving the financial incentives for investment, not direct federal participation.

The Administration will therefore be expected to drop support for any future co-operative generic technology centres, even though specific legislation setting up these centres was approved by Congress last year, and \$5 million allocated to their support (*Nature* 286, 195, 1980). From the three centres initially proposed, only one, in Detroit, is likely to survive as plans were agreed before the election, although even this is uncertain since it will depend on the centre's ability to raise matching funds from industry.

Other projects previously under development in the Commerce Department and now expected to be phased out include proposals to establish state-based Corporations for Innovation Development to help entrepreneurs gain access to investment capital, and the new Office of Technology Strategy and Evaluation.

It is also rumoured that the new Administration may not seek a successor to Dr Baruch, who as Assistant Secretary for Productivity Technology and Innovation was responsible for science and technology programmes within the department. This would be opposed by Congress, which spent much time last year discussing the Carter Administration's initiatives, and supported their general thrust.

Another of Mr Carter's initiatives which will be overruled is the Co-operative Automotive Research Project (CARP), originally proposed by Transportation Secretary Mr Brock Adams as a means of

JET upgraded

The extension of JET to full design performance would involve increased magnetic confinement fields, additional heating of the plasma, facilities for storing and handling tritium, and remote handling equipment for experiments. Specifically, the peak toroidal field coil power would be increased from 250 MW to 380 MW, total magnetic field at plasma centre 27.7 kG to 34.5 kG, plasma current from 3.8 MA to 4.8 MA, and additional heating power from 4–10 MW to 25 MW.

Robert Walgate

generating joint government and industry support for long-term research, and already approved by Congress with a budget of \$12 million for the first year. In his budget proposals of two weeks ago, Mr Reagan said that "federal financing of long-term research to benefit a particular industry is an inappropriate allocation of federal funds." He is proposing to rescind the full amount appropriated by Congress for the current year, and to terminate the whole programme.

At the National Science Foundation, schemes for encouraging greater technological innovation in small businesses, and for forging closer links between industry and universities, will remain in force, but will not get the substantial increase in funding that the Carter Administration had promised as part of its innovation package.

Support for small industries innovation, for example, was to have been almost doubled, from \$7.5 million to \$14.5 million next year, following its earlier success, but will now be cut back. So too will increased funds for engineering education, but the proposed 20 per cent increase for engineering research is likely to remain.

In contrast with the cuts being proposed in measures which would increase federal involvement in the innovation process, other steps initiated by Mr Carter to reduce the federal role have been warmly endorsed and built upon by the new Administration.

Efforts to reduce the burden of health, safety and environmental regulations, for example, have already been expanded. As expected, Mr Reagan has proposed, along with his budget reductions, a set of regulatory reforms which would submit all new and existing regulation to strict cost-benefit analysis.

Similarly, additional patent reform legislation has already been introduced into the new session of Congress which would expand on Mr Carter's patent reform bill giving universities and small businesses patent rights on federally funded research.

David Dickson

UK research councils Allen accused

The UK Science Research Council is being hauled over the coals for sloppy bookkeeping. Its chairman, Sir Geoffrey Allen, is to appear before the Public Accounts Committee of the House of Commons on 17 March to explain irregularities in the council's funding in the financial year 1979-80. Details are hard to obtain because the council, usually frank, is saying nothing for fear of offending parliamentary privilege.

Part of the problem stems from the cash bonus that Mrs Shirley Williams, then Secretary of State for Education and Science, obtained for the research councils in 1978. The council's share was £33 million over the four financial years

1979-83. A circular was thereupon sent around universities asking for applications to the council to replace worn-out equipment. Truck-loads of applications followed, worth £37 million, of which the council awarded £7.5 million. In the event, a change of government followed, funds were cut and the council received only £5 million.

Another question mark hangs over the university grants current in March 1980, which represented an increase in value of £31 million (34 per cent) over the previous year. There appears to be no indication that this large increase was planned. The exact amount of overspending remains unclear, but much of the money was spent by the Science Board (responsible for such topics as physics of solids and liquids, chemistry and biology). The new Spallation Neutron Source at the Rutherford-Appleton Laboratory at Chilton seems to have been a principal beneficiary.

Other misdemeanours are procedurally more serious. The Auditor General, Sir Douglas Henley, has already complained that the postponement of certain payments into the 1980-81 financial year, as part of an attempt to alleviate the financial deficit, contravenes government regulations. The Public Accounts Committee will also, no doubt, be asking about the council's calculation that the capital value of a site near Slough, yet to be vacated, could be regarded as a part of the income for 1979-80.

The council appears to have been the victim of government financial vagaries combined with inflexible accounting procedures. But the extent to which the problems are also self-generated will not be clear until 17 March.

Sir Geoffrey Allen, formerly the council's accounting officer, came to the end of his spell as chairman in October. The name of his successor is expected to be announced within a few weeks.

Philip Campbell

Soviet research

More home growth

The new Soviet Five-Year Plan calls for all branches of the economy to be brought up to the "most up-to-date levels of science and technology." Just how to do this is clearly causing the Soviet leadership considerable anxiety. At the Twenty-Sixth Congress of the Communist Party of the Soviet Union last week, Mr Brezhnev called on the whole scientific establishment to reassess the research and development basis of Soviet industry and to propose ways of regrouping the "scientific forces".

Not only the Academy of Sciences and the State Committee for Science and Technology should take part in this audit, said Mr Brezhnev, but also the science-based industries, including defence. Since Soviet military research is organized quite

separately from the civil sector, this last proposal suggests genuine concern, not simply congress window-dressing.

Mr Brezhnev singled out a number of fields of technology where "impermissible sluggishness" had led to delays in implementing "promising developments" — the continuous casting of steel, powder metallurgy, custom-built DC transmission lines and high-strength artificial fibres. Falling behind foreign competitors, he said, leads to massive expenditure of foreign currency for equipment and technology which the Soviet Union could have produced at home. Soviet potential technological self-sufficiency has been a feature of propaganda speeches since the January 1979 United States embargo. Mr Brezhnev's speech, however, referred rather to one of the major concerns of Soviet research policy: why is there often so long a gap between obtaining a new result and implementing it in production?

Mr Brezhnev suggested two possible lines of reorganization, which appear mutually contradictory. On the one hand, he stressed the Central Committee's support for an increased responsibility for the Soviet Academy of Sciences, and argued a "flexible and mobile" organization of research that would not tolerate "fruitless laboratories and institutes", but would respond "attentively" to the needs of scientists for equipment, instruments and pilot plant facilities. Taken in isolation, these remarks suggest more scope for serendipity and the capacity to switch rapidly from one line of research to a more promising alternative.

Mr Brezhnev went on to say, however, that the major sciences (including basic research) should concentrate more on solving "key national economic questions" and "discoveries capable of making genuinely revolutionary changes in production". The formulation of these tasks, he said, is the task of the central planning bodies and the State Committee for Science and Technology. The exact spheres of competence of the Academy and State Committee are frequently difficult to define, and Mr Brezhnev's speech does not make the issue easier. The previous Congress (1976) had made the Academy responsible for coordinating all science throughout the country, and although, to judge from the report to Congress of Dr Anatoli P. Aleksandrov, the Academy's president, much still remains to be done, there is no suggestion that the task should be taken out of the Academy's hands.

Dr Aleksandrov's report, moreover, reviewed a wide range of recent achievements, from particle physics and cosmology to the utilization of Estonian shales and the need to develop coal liquefaction and gasification techniques. Academy scientists, he said, have made notable advances in thermonuclear fusion, and in prolonging the life of agricultural machinery.

Discussion of future plans, at a Congress

of this type, is always to a large extent, orchestrated. A clear example was the intervention of Sharaf Rashidov, Party chief of the Uzbek SSR. Mr Brezhnev had called for a new food production programme. Mr Rashidov suggested that a "tremendous contribution" could be made by diverting the northern rivers to the Volga Basin and the Siberian rivers to the steppes of Kazakhstan and Central Asia. Plans for such proposals, however, are of course already well under way, and a considerable engineering effort earmarked for them in the new Five-Year Plan.

Dr Aleksandrov, however, had a lightertouch. One of his major proposals began as an apparent piece of by-play with Mr Brezhnev and the Leningrad Party chief Grigori Romanov, about the development of an anti-influenza spray which Mr Brezhnev had requested at the last Congress. (The very fact of such jesting is, in itself, a suggestion that the Academy's auspices are exceptionally favourable at present.) The successes which Dr Aleksandrov reported — 10 million doses of the spray produced in 1980, 25 million scheduled for 1981, and the incidence of influenza in Leningrad cut by two-thirds, were simply a build-up. Soviet microbiological production, suggested Dr Aleksandrov, should be reorganized under a special Ministry of the Biotechnical Industry. **Vera Rich**

German nuclear power

Closures and delays

Four German nuclear power stations, representing a third of the country's 10,000 MW of nuclear power, are to close for at least a year for the complete replacement of their primary cooling circuits. The Bonn Nuclear Safety Commission ordered the closures after deliberating for three years over this issue, involving General Electric Company-designed boiling water reactors (BWRs) constructed under licence between 1973 and 1979.

Small leaks and cracks had developed in the primary circuit of these reactors which carry cooling water through the nuclear cores. The cracks appeared to be spreading by stress corrosion (combined effects of mechanical and thermal stress and chemical corrosion). The commission has concluded that uneven quality, poor choice of materials, and a thin gauge of steel for

the pipework were to blame, and that the circuits must be replaced to avoid the danger of a potentially catastrophic cooling failure. The cost of replacement has been estimated at DM 1,200 million (£250 million), with an equal sum for the loss of electricity sales. However, the four responsible utility companies have decided not to close down all the reactors simultaneously. One is already being refitted, and should be on-power again by the summer, and the others will be dealt with in sequence.

The boiling water reactors have not been a success, and Germany is now concentrating on the Westinghouse pressurized water reactor design, originally licensed by Siemens. Siemens and the constructors, AEG, are now merged in the company Kraftwerk-Union, and the excessive cost-trimming which may have led to the BWR problems is not expected again.

Meanwhile, the federal government's nuclear policy continues to be threatened by opposition from the left wing of the Social Democratic Party, particularly in Hamburg where construction of the Broksdorf reactor has been brought to a halt by demonstrations. Over the weekend, 50,000 people demonstrated at Broksdorf, and 127 policemen were injured. According to some observers, many of the demonstrators were not local, but represented national left-Social Democratic, Communist and Green Party groups. Broksdorf thus appears to have become a symbol of left-right conflict in Germany. **Robert Walgate**

Human growth hormone

Pituitary slump

Britain's supply of pituitary glands, from which human growth hormone (HGH) is extracted to treat 600 British children who do not produce it normally, has fallen by at least a third in the past year — to a level too low to keep up with demand. This has happened since responsibility for HGH extraction shifted, in June last year, from two Medical Research Council-funded researchers to the Department of Health and Social Security (DHSS). Bureaucracy has strangled their collection procedure, Dr Philip Lowry, one of the researchers, said last week. The collection rate has now stabilized at around 42,000 glands a year, DHSS claims, accounting for almost all patients who end up in a National Health Service mortuary — but the supply from public mortuaries (where victims of violence and accident are given postmortem examinations) has almost dried up. In the last year of the Medical Research Council scheme, 70,000 pituitaries were collected, said Dr Lowry.

Extraction of the hormone is now done at the Centre for Applied Microbiological Research at Porton Down, which has become increasingly involved in health

service functions since it was transferred to the Public Health Laboratory Service from the Ministry of Defence two years ago. Of two processes, depending on the form in which pituitaries are collected, Porton chose the older, using acetone-dried glands. This yields around 8 clinical units of HGH per gland as opposed to 15 by the newer frozen-gland system (which has been adopted in New Zealand and Canada), but collection is easier: the extracted gland needs only to be popped into a bottle of acetone, instead of a container packed around with dry ice. However, says Dr Lowry, who developed the frozen-gland process, clinical trials over the past year have shown that the HGH produced from frozen glands is the more pure, provoking no immune reactions — unlike acetone-dried HGH, which often contains modified and agglomerated versions of the protein, and can be rejected by the child.

The 336,000 units of HGH made available by the Porton process each year in Britain are enough to supply only 430 of the 600 children with a full dose of 15 units a week. So far, Porton has been relying on an excess of material collected through the old system, but soon the reduced supply of glands must have an effect on the supply of HGH. It may then be necessary to buy HGH from foreign producers, where market pressures have pushed up the price to 50 times the cost of home-produced hormone. Porton is also working on producing HGH from *Escherichia coli*, genetically engineered to excrete the hormone, provided by the Swiss firm Kabi-Vitrum in association with the California-based genetic engineering company, Genentech, but production is not likely to be significant for a year, at least until clinical trials, now under way at Great Ormond Street Children's Hospital on Genentech-produced hormone, are complete. Moreover, the genetically-engineered hormone is slightly modified: it has an extra amino acid (methionine) tacked on one end, and there is a possibility that, like agglomerated HGH, it may provoke an immune response.

Meanwhile, the DHSS is trying desperately to restore supplies of pituitaries from public mortuaries. But the department is facing numerous legal and personal obstacles. Legally the coroner must consent, and the relatives give written permission. But a death involving a public post mortem is usually an unexpected one, relatives are shocked, and the case does not seem so compelling as for a kidney or a heart. The DHSS recently sent 200 letters and made as many telephone calls to coroners and pathologists requesting their urgent cooperation — but the department has little hope that the supply will rise much in consequence.

A few months ago, supply problems were even more desperate, when hospitals too were failing to provide the glands. This may have been because morticians, in March 1980, negotiated a new contract

In addition to the four boiling water reactors (producing 3,300 MW), Germany has another six power reactors producing 6,600 MW. Another nine, to produce around 10,000 MW, are under construction, including the SNR fast breeder research reactor, and a high temperature reactor near Cologne. Construction is halted at two further sites (Broksdorf and Wyhl because of local and political opposition).

under which their wages were increased to cover extra duties — which included the extraction of pituitaries. Previously they had been paid on a "piece rate" — 20 pence a gland — so they now had no incentive for their unpleasant task. Pressure from hospital authorities on the workers, however, brought a response. Public mortuary workers also now receive nothing where they received 20 pence per gland before, so this may also be a factor in the HGH supply shortage. Negotiations are under way, says DHSS, on how these workers might be compensated by the area hospital authorities for providing the glands.

Some officials are suggesting that the real cause for the fall in supply is the "regularization" of collection procedures, compared with a little bit of horse-trading before. In the past, acetone sample bottles have been sent through the post, this is strictly illegal (acetone being inflammable) and has now been stopped. Also, extraction is now being done under full containment because of the danger of the propagation of slow virus from an infected gland (the pituitary being nervous tissue).

And were relatives always informed? Dr Lowry claims they were, and that he thus obtained more than half the pituitaries available from public mortuaries in London. The problems he lays squarely at the door of the DHSS. **Robert Walgate**

University of London Storm ahead

The reorganization of the University of London promises to be a more turbulent process than was foreseen a year ago. The plan that the Swinnerton-Dyer committee should make its final recommendations by the end of this year has been foreshortened by the mounting sense of urgency within the university. The committee has now agreed, with some misgivings, to produce its recommendations before the beginning of the next academic year. Meanwhile, the basis of the committee's interim report has been questioned by a document now circulating within the university, and likely to colour the arguments put by the Association of University Teachers to the committee at a meeting next week.

That document questions the arithmetic leading to the conclusion that the university's income will fall by between £15 and £20 million by the end of this decade. This is held to involve the unwarrantable assumption that the 3½ per cent cut in university finances announced last December will be permanent, the unnecessary assumption that overseas students attending London colleges will be charged only the minimum tuition fees decreed by the government and inconsistent assumptions about the basis used for calculating the loss of income if the numbers of overseas students decline. Briefly, the committee's calculation for the university as a whole is said to exclude the allowance made by the University Grants

Committee for part-time students.

This and other documents will form the basis on which the union will draw up a formal reply to the Swinnerton-Dyer committee's interim report, probably before the end of the month. Apart from arithmetical arguments, the union is concerned at the negative tone of the interim report, which is said to have neglected the university's advantages, in particular its attractions for overseas and part-time students and the relatively favourable age structure of its academic staff.

This view is likely to be echoed by the responses from individual colleges in the university, expected by Easter. Another view gaining ground is that the committee's interim report, with its comprehensive analysis of how the university spends its money, is sufficient, and that it should be for the university rather than the committee now to say what should be done.

The pitfalls for such committees are nicely illustrated by the troubles which have befallen the Flowers report on the organization of the medical schools in the University of London. After the rejection of the Flowers recommendations by the senate of the university last October, the

Joint Planning Board set up a working party to make a more detailed analysis of the costs of medical education in the University of London. The report, based on an analysis by a firm of professional accountants, suggests that the cost savings obtainable by closing various clinical and preclinical schools are not those forecast by the Flowers committee, and that none of the options so far discussed would yield as large a saving on the total cost of £32.8 million as a modest change in the staff-student ratio.

Chemical and nuclear weapons Scientists speak out

Two groups are being set up to give British scientists a stronger voice on the development of weapons of mass destruction. On 26 February, the Russell Committee Against Chemical Weapons, a branch of the Bertrand Russell Peace Foundation, launched a campaign against what is seen as a chemical arms race. And at the end of this month, a conference is being held at the Open University, Milton Keynes, to inaugurate Scientists Against Nuclear Arms, a group committed to providing scientific information on nuclear weapons.

Scientists Against Nuclear Arms says that it has already received several hundred letters expressing interest and support. And 19 signatures were attached to the appeal for names of those opposed to chemical weapons when it was launched last week. One notable signatory, Dr Frederick Sanger, not distinguished as a signer of petitions, is particularly concerned with the possible use of new techniques in biology to develop more sophisticated chemical weapons and with the distortion of research priorities that could result from a chemical arms race. These issues, and uncertainties about the British government's future policy, have prompted the appeal, which asks scientists not to take part in research on chemical weapons and the British government not to stockpile them.

Despite repeated assurances that there are no plans to store or manufacture chemical weapons in Britain, the appeal's organizers fear that the British government's policy may change in response to pressure from the new Reagan Administration to house stocks of new binary chemical weapons on European soil. Last year, the United States announced that it was building a plant to manufacture binary weapons.

Believing the present to be a time of uncertainty, the appeal's organizers say there is an urgent need for public and parliamentary debate. Scientists have an important role, they argue, because the development of chemical weapons could be influenced more directly than that of any other type of weapon by research in non-classified laboratories. **Judy Redfearn**

Soviet activists honoured

Three leading Soviet human rights activists have recently been honoured by Western scientific bodies.

The most prestigious award, foreign associateship of the French Academy of Sciences, was conferred on Dr Andrei Sakharov on 16 February. This, however, according to M. Paul Germain, one of the two secretaries of the academy, was in no way a political act, but was simply in recognition of the great importance of Sakharov's work in various fields of physics.

The Catholic University of Louvain in Belgium however, took a double view. When challenged by the Soviet embassy in Brussels that they should not award an honorary doctorate to Viktor Brailovskii, the Moscow refusenik cyberneticist, on the grounds that he was a "criminal and a prisoner" the rector replied that they were conferring the award in recognition of his contribution to the field of science and for maintaining the weekly seminars (for refusenik scientists).

Finally, the Royal College of Psychiatrists in London, in conferring an honorary fellowship on Dr Semeon Gluzman, did so specifically in connection with his work for human rights, for Dr Gluzman's chief contribution to psychiatry was the co-authorship (with Vladimir Bukovskii) of the *samizdat* "Handbook of Psychiatry for Dissidents" (a manual on how to resist psychological and pharmacological pressure), for which he is now serving a three-year sentence in Siberia following seven years in a labour camp.

CORRESPONDENCE

Data bank

SIR — The news item "Nucleotide sequences Too many banks" (*Nature* 15 January, p 112) implied that we do not yet have a nucleic acid sequence data bank in fact our data base system contains the largest collection of nucleotide sequences in the world, updated on a regular basis and accessed through a sophisticated computer retrieval system

The data bank was first made available to the scientific community by telephone link to our computer on 15 September 1980, when the information base contained over 227,000 residues of DNA and RNA sequences published before 1 August 1980. During a three month demonstration project, the collection was updated monthly with material published mainly between August and October. The final total on 12 December 1980 was 350,000 residues

We are continuing to update the data base and to make it available on a subscription basis. We are now processing more than 40,000 residues a month, and a book is being prepared which will include the information in the collection and describe the retrieval system and its uses

The data bank is supported by the Institute for General Medical Science of the US National Institutes of Health, the National Aeronautics and Space Administration, and contributions and subscriptions from foundations and industrial concerns. The expense involved in maintaining a complete, critically reviewed and up-to-date data base is too great for it to be operated on a voluntary basis, so other means of support must be sought. Our data base is available on a subscription basis (for the use of the subscriber in his own institution, not for redistribution). The collection will be made available on this basis to the European Molecular Biology Organization for use at Heidelberg, under the direction of Dr Greg Hamm, so that the interests of the European non-commercial research community can be considered

The cost of such a data base is usually justified by the resulting benefit to a research area. If all those organizations that could benefit from a nucleotide sequence data base were to contribute some support, there would be sufficient money to establish a continuing project

M O DAYHOFF, R M SCHWARTZ, H R CHEN,
L T HUNT, W C BARKER, B C ORCUTT

National Biomedical Research Foundation,
Georgetown University Medical Center,
Washington DC, USA

Antipodean 2,4,5-T

SIR — There is continuing concern in Australia regarding the possible deleterious effects of the herbicide 2,4,5-trichlorophenoxyacetic acid (2,4,5-T) on animal life, especially human beings. The question of a link between 2,4,5-T usage and congenital abnormalities has been raised on several occasions¹⁻³. Concerns of this nature were major reasons for withdrawing 2,4,5-T-based defoliants (such as Agent Orange) from use in the Vietnam war during the early 1970s. Many batches of defoliant

contained high quantities of the extremely toxic contaminant 2,3,7,8-tetrachloro-cyclo-dibenzo-*p*-dioxin, suspected of being a teratogen. International sale of the unused defoliant was apparently considered by the US government⁴ and some of this defoliant may have been imported into Australia⁵.

On 18 August 1971 the Australian government initiated enquiries under the provisions of the Customs Tariff (Dumping and Subsidies) Act regarding the importation of "salts of 2,4,5-trichlorophenol". Two years later, dumping duties were imposed on goods based on the esters and salts of 2,4,5-T, *retrospective* to 4 January 1971. It seems likely that large quantities of such goods were bought very cheaply overseas and sold on the Australian market in the late 1960s or early 1970s.

The chemicals 2,4,5-T and 2,4-dichloro-phenoxyacetic acid which make up the defoliant Agent Orange would normally be classified as pesticides in trade statistics. However, they could be given a chemical classification and assigned the 1972-3 SITC code number 512 28 09 ("other phenol derivatives, halogenated etc"). In the financial years 1969-70 and 1970-71 over 300 tonnes of chemicals with this SITC code are recorded in Australian trade statistics as imports from Singapore (the "country of production"). None of this material is listed as an export or re-export of Singapore in that country's *External Trade Statistics*, and there have been no subsequent imports from Singapore. (A small import from Singapore is recorded in 1968-69, but there is nothing before this.) The Australian Bureau of Customs has declined to be of any assistance in determining the true nature or origins of these goods.

With the help of the Australian Bureau of Statistics we have ascertained that in 1970-71, 290,000 lb of these imports entered Australia through Queensland, and another 22,400 lb through Western Australia. The *Australian Chemicals Guide (Industry Study)* shows a very limited number of companies involved in the processing of halogenated phenols. With the assistance of the *Guide* and the reference *Who owns Whom* it is possible to obtain a very short list of potential importers.

The Australian government has access to the records which would permit a thorough investigation of this matter. The results of such an investigation could have a vital influence on a multi-million dollar epidemiological study at present being undertaken by the Australian government on the effects of defoliants on Vietnam veterans. The study envisages a comparison of veterans exposed to defoliants with those "unexposed" veterans who did not serve in Vietnam.

Further details of this matter appear in a forthcoming paper⁶.

PETER HALL*, BEN SELINGER*

BARBARA FIELD†, CHARLES KERR†

*The Australian National University,
Canberra, and

†The Commonwealth Institute of Health,
Sydney, Australia

- 1 Report of the Consultative Council on Congenital Abnormalities in the Yarram District (Government of Victoria, Melbourne, 1978)
- 2 Field, B & Kerr, C *Lancet* ii, 1341 (1976)
- 3 Hall, P & Selinger, B *Chem Aust* 47, 420 (1980)
- 4 Shapley, D *Science* 180, 43 (1973)
- 5 Dux, F & Young, P J *Agent Orange — The Bitter Harvest*, 38-41 (Hodder and Stoughton, Sydney, 1980)
- 6 Hall, P & Selinger, B *Chem Aust* (April 1981)

Cell contamination

SIR — The editorial, "Responsibility for trust in research" (*Nature* 22 January 1981, p 211) was excellent. Our group sends various human cell lines to laboratories all over the world. We caution scientists about the dangers of cell-cell contamination and if we read an article with strange conclusions that involved one of our many hundreds of cell lines, we offer to re-authenticate its identity without cost. Often the cell line in question has been passed from one laboratory to another. We have been discouraged to detect publications in which the scientists have continued to use a contaminated cell line even after the error has been pointed out to them. Biochemists in particular seem loath to seek biological advice. One chemist replied that the kind of cells that he used wasn't really important.

Actually there are other serious contaminations which affect research with cells. Most cell lines are cultured in media supplemented with antibiotics and often the cultures are contaminated with slow-growing bacteria and mycoplasma. It is difficult to detect minimal numbers of these organisms. Electron microscopy of a pellet of centrifuged media may reveal them.

We do not agree that the responsibility for monitoring the validity of research rests with the department head. Departments are too complex and even the division heads may not have the competence to evaluate a specific "bench" operation. The scientist himself and his associates in a "specialty" must exert the self discipline of avoiding, acknowledging and correcting errors.

As the use of enzyme profiles and new techniques of cytogenetics and the detection of "specific" antigens and cell products evolve, it will be easier to detect cell-cell contamination — perhaps even hybridization. Our methods for authenticating cells are still relatively primitive. The establishment of authorities to referee cell-cell contamination is premature. The donor and recipient of a cell line share the responsibility of the authenticity of the cell line.

Identification of a cell line should include (1) data about the original collection of tissue and pathological diagnosis, (2) serial numerical records of all subcultures, (3) electron microscopy — preferably of both the tissue of origin and the cultured cells, using special stains if critical, (4) isoenzyme profiles, (5) karyotype analysis of chromosome morphology after banding, (6) cell products, (7) studies of growth in immunosuppressed animals such as nude mice, and (8) growth characteristics in various media.

The errors associated with cell lines have probably had an infinitesimal effect on the fabric of scientific progress in comparison with other ridiculous and shameful pronouncements. For example, governmental carcinogens, the clinical effectiveness of interferon, and numberless epidemiological studies such as telephone surveys of complex disease states.

GEORGE E MOORE
LINDA K WOODS

Department of Surgery,
Denver General Hospital,
Denver, Colorado, USA

NEWS AND VIEWS

Carbon dioxide and climate: ice and ocean

from Starley L. Thompson and Stephen H. Schneider

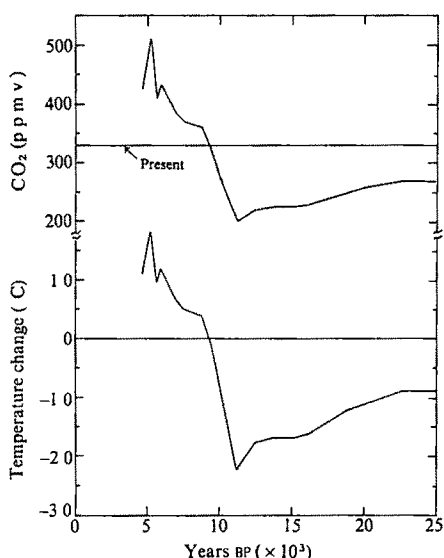
THE presence of carbon dioxide in the atmosphere has long been known to affect climate through the so-called 'greenhouse' effect. The present rate of growth of atmospheric CO₂ concentration due to fossil fuel burning will result in a doubling of the CO₂ level sometime in the next century and this may produce a global surface temperature increase of between 1.5 and 4°C. Since this temperature change is as large or larger than any natural change in the past few millennia, the increasing interest in the connection between CO₂ and climate is not surprising. Recently, new ground has been broken in two relatively unexplored areas — the measurement of natural CO₂ variations on long time scales and the role of the ocean in modulating climatic response to the anthropogenic CO₂ increase.

The absence of firm evidence of past CO₂ variations has been a troublesome gap in our knowledge. A correlation between levels of CO₂ and palaeoclimatic temperature data would constitute the first clear empirical evidence of CO₂-induced climatic change.

Two recent papers have shown that past levels of atmospheric CO₂ concentration can be reconstructed from the CO₂ trapped in large ice sheets. Berner *et al.*¹ and Delmas *et al.*² have analysed trapped air in polar ice cores and concluded that large variations of atmospheric CO₂ occurred during the past 30,000 years. Previous attempts to analyse the total CO₂ content of the ice cores failed, but new extraction techniques increase the researchers' confidence that the measured CO₂ variations may reflect accurately the CO₂ variations in the atmosphere.

Two features stand out in the accompanying figure from Berner *et al.*¹ First, their data show a CO₂ concentration of as little as 200 p.p.m. by volume 10–30 years ago compared with a present-day value of around 340 p.p.m. Second, there appears to have been more atmospheric CO₂ in the mid-Holocene (about 5,000 years ago) than at present. Estimates from

current climatic models of the global surface air temperature change caused by such CO₂ variations are also shown. Uncertainties in present climatic theory suggest that these temperature estimates should be within a factor of two of the actual CO₂-induced temperature changes — assuming, of course, that these preliminary measures stand up to further analysis. Note that the estimated global temperature changes range from +1.8°C to –2.2°C. These represent a large fraction of the temperature range deduced from other palaeoclimatic evidence for the period.



Graph showing the relationship between the atmospheric CO₂ level reconstructed from Camp Century, North Greenland, core measurements (upper curve) and the corresponding global change in surface air temperature estimated from current climatic models (lower curve) (Taken from ref. 1).

Should these large CO₂ variations prove to be reasonably accurate, the implications for palaeoclimatic theory would be enormous. One of the most difficult problems in palaeoclimatic modelling has been the determination of the feedback processes which must exist to cause the large-amplitude glacial to interglacial transitions.³ A feedback including natural CO₂ variations like those in the figure could explain as much as half the surface temperature amplitude of the ice age/inter-

glacial climate changes. However, it does not appear that the CO₂ variations initiate these large climate changes as the minimum and maximum CO₂ concentrations do not appear to lead the respective glacial and interglacial extremes of 18,000 and 6–8,000 years ago. Therefore, CO₂ changes may act only to amplify climate changes that are initiated by other means.

In addition to improving our understanding of past climates, any evidence that CO₂ levels were higher during the mid-Holocene than today may give more credence to the idea of using this climatic 'optimum' period — when surface temperatures were a few degrees warmer than today — as an analogue for the climatic change which may be induced by a future anthropogenic CO₂ increase.⁴

A new area of research on the CO₂ problem has been recently explored by Hoffert *et al.*⁵ Previously, most climatic modelling studies examining the effect of increased CO₂ have considered the models' response in a state of climatic equilibrium. For example, a researcher would perturb a mathematical model of climate with a fixed, time-independent doubling of the CO₂ concentration and allow the model climate to reach a new equilibrium. This is a reasonable approach to take when testing the various models since it affords an order-of-magnitude estimate of climatic response for a fixed forcing (for example, when the CO₂ concentration has doubled), while meeting the constraints of a limited computing budget. However, the actual CO₂ increase is a continuous rise, not an instantaneous step increase. In addition, because of its large heat capacity, the ocean acts as a thermal buffer by delaying the response of the atmospheric temperature — an effect that has been neglected or undervalued in most equilibrium studies.

In order to simulate adequately the climatic response to the actual time-dependent CO₂ increase one must account for oceanic heat capacity, which means that one must determine how the ocean mixes the excess CO₂ heating at its surface. The upper layer of the ocean is mechanically well mixed by surface winds to a depth of approximately 50 metres. Including the heat capacity of the mixed layer alone would delay the global temperature response to the CO₂ increase.

Starley L. Thompson is in the Department of Atmospheric Sciences, University of Washington and Stephen H. Schneider in the National Center for Atmospheric Research, Boulder, Colorado

by 5 to 10 years^{6,7} But the mixed layer is only a small fraction of the ocean volume. Mixing occurs throughout the depth of the ocean, although at much slower rates than within the mixed layer. When Hoffert *et al* simulated the heat capacity of the entire ocean in an ocean model with vertical diffusion and specified upwelling, they found that the response of the global temperature to the CO₂ increase would be delayed by 10 to 20 years by the year 2,000. Even longer lags would occur in the next century.

The effect of such a global time lag will have to be considered when examining

future climatic data for the predicted CO₂ warming. We may not be able to extract a detectable CO₂ 'signal' from the climatic 'noise' (ref.8) until several decades after the date estimated from equilibrium calculations. Thus the absence of unambiguous warming during the remainder of this century would not serve to disprove the validity of our current equilibrium estimates of the effect of CO₂ on global temperature.

However, the determination of the global temperature time lag cannot be divorced from the complex problem of modelling the regional variations in oceanic mixing. As much of the communication between the deep ocean and surface waters occurs in downwelling polar seas, Hoffert *et al* concluded that the as yet poorly understood response of the polar oceans to the CO₂ increase may be very important in determining the global time lag.

A related idea (discussed by Schneider and Thompson⁹) is that variations in oceanic mixing which create a geographical distribution of effective oceanic heat capacity may also serve to produce thermal time lags with substantial regional

variations. The significance of this goes beyond the added difficulty in unambiguous early detection of the global CO₂ warming. Different thermal lags for different regions implies that anomalies in surface temperature gradients, and any associated regional climatic anomalies, may with time become substantially different from those computed in an equilibrium simulation. Since it is the potential regional climate changes which are of the most economic concern, we feel that improving our understanding of the nature of the transient climatic response should be a major priority for future CO₂-climate research. This means more coupled atmosphere-ocean-cryosphere modelling studies.

Finally, we should point out that the problem of the transient climatic response is just one of a set of interdisciplinary problems. How much fossil fuel will be burned, the global carbon cycle, the global and regional equilibrium climatic response, the transient response, the physical and societal impacts of climatic changes, and possible policy responses to prevent or adapt to CO₂ effects are all important aspects of the 'CO₂ problem'. □

- 1 Berner, W., Oeschger, H. & Stauffer, B. *Radiocarbon* 22, 227 (1980).
- 2 Delmas, R. J., Ascencio, J.-M. & Legrand, M. *Nature* 284, 155 (1980).
- 3 Schneider, S.H. & Thompson, S.L. *Quater. Res.* 12, 188 (1979).
- 4 Kellogg, W.W. & Schwart, R. *Climate Change and Society* Appendix C (Westview Press, Boulder, Colorado, 1981).
- 5 Hoffert, M.I., Callegari, A.J. & Hsieh, C.-T. *J. geophys. Res.* 85, 6667 (1980).
- 6 Thompson, S.L. & Schneider, S.H. *J. geophys. Res.* 84, 2401 (1979).
- 7 Ad Hoc Study Group. *Carbon Dioxide and Climate: A Scientific Assessment* (US National Academy of Sciences, Washington DC 1979).
- 8 Madden, R. & Ramanathan, V. *Science* 209, 763 (1980).
- 9 Schneider, S.H. & Thompson, S.L. *J. geophys. Res.* 86 in the press.

Active galaxies: dragon hunting in Arizona

from Martin Elvis and Andrew S. Wilson

PLOTS of the electromagnetic spectrum of active galaxies covering the ten decades of frequency from radio waves to X rays were described at a recent workshop* and show just how rapidly this field is advancing. The successes of the radio astronomers' Very Large Array (VLA) and the HEAO X-ray observatories enabled radio and X-ray astronomers to meet their optical colleagues on nearly equal terms for the first time. The sensitivity and angular resolution of the new instruments is bringing some order into the classification of active galaxies — the old diversity now appears to have been largely the result of observations limited to only the extreme examples of particular characteristics.

The meeting was restricted to active galaxies and conspicuously avoided their close relatives, quasars. As a result there was an emphasis on objects in which resolved structure could be seen. The most striking examples of this came from the detection of over 50 jets by the VLA (reviewed by Bridle, VLA, New Mexico). These long, collimated structures point from the nuclei of bright radio galaxies to their giant radio lobes. They needed a dynamic range approaching 1,000:1 to be discovered close to the bright nuclear sources but, now that this has been achieved, observations of such jets have

become common, almost ubiquitous. Over 60 per cent of the nearby ($z < 0.05$) 3CR radio galaxies have them. The jets are seen on all size scales from 1 parsec to 1 Megaparsec. Often they are one-sided, perhaps as a result of relativistic beaming. In all likelihood we are seeing the beams of relativistic particles which Rees proposed ten years ago would power the giant radio lobes. Polarization measurements show that the magnetic field is often parallel to the jet near the nucleus but turns at right angles to the jet further out (although sometimes leaving a parallel region at the edge). There is a suggestion that the parallel region is longer for higher radio core luminosities. If such a correlation is borne out it will provide useful information for the beam instability models discussed by Blandford (Caltech), Laing (NRAO, Virginia) and van Bruegel (Leiden).

It is necessary that the jets are confined in some way. The simplest possibility is that the galaxy and the jet are surrounded by hot gas. This should be visible by X-ray emission and observations have been made from jets in the three classic sources: M87, Cen A and 3C273 (Schreier, Harvard-Smithsonian Center for Astrophysics). However, the detailed correspondence of the X-ray, optical, and radio features suggests strongly that the X rays are a continuation of the jets synchrotron spectrum. For other 3CR radio galaxies initial results of the *Einstein* survey

(Fabbiano, Harvard-Smithsonian Center for Astrophysics) require equipartition fields within the jets of the order 10^6 – 10^7 gauss. If these radio jets are thermally confined the expected X-ray emission from the confining gas is just below the current detection threshold.

Miniature double sources have been found by Wilson (University of Maryland) in five Seyfert 1 galaxies. Type 1 Seyferts are the archetypal 'radio-quiet' active galaxies so this discovery is particularly interesting. It demonstrates that the same linear symmetry we see so clearly in the radio-loud case applies to this other, much more numerous, class. Since these Seyfert 1 galaxies have clear spiral structures it is simple to see whether the direction of radio symmetry aligns with the rotation axis. Such convenient alignment does not appear to occur.

Before leaving jets the extragalactic astronomers were presented with a 'working scale model' of their problem. Hjellming (VLA, New Mexico) presented a beautiful series of VLA observations of SS433 showing a complex series of changes with time. He was able to explain these with a simple, constant velocity, ballistic blob model. Whilst this is surely too simple for the extragalactic case it must have some

Martin Elvis is in the Center for Astrophysics, Cambridge, Massachusetts and Andrew S. Wilson is in the Astronomy Program, University of Maryland.

*A workshop on active galaxies was held in Tucson, 19–23 January 1981 and was sponsored by the Kitt Peak National Observatory.

suggestive power for those trying to understand the larger scale phenomenon

The other three days of the meeting were used to examine and compare observations of the compact nuclei themselves. Weedman (Pennsylvania State University) presented a useful criterion for distinguishing the emission line galaxies that need a non-thermal ionizing continuum from those that can be explained solely by bursts of star formation. He found that for emission line widths greater than 250 km s^{-1} at least 90 per cent of galaxies needed, as he said, 'a Dragon in the middle' — a general term for any source of a non-thermal continuum. For narrower line widths no dragon is needed.

Far from being discarded for their lack of activity narrow line 'starburst' nuclei emerged as a new area of study with radio (Gisler, NRAO, Virginia), $\text{H}\alpha$ emission line (Heckman, NRAO, Virginia) and X-ray (Fabbiano, Harvard-Smithsonian Center for Astrophysics) images providing new inputs. Why some galaxies undergo intense star formation periods is as little understood as why other galaxies have non-thermal active nuclei.

Further clarification of what constitutes an active nucleus came from comparing the X-ray (Mushotzky, Goddard Space Flight Center) and optical (Huchra, Harvard-Smithsonian Center for Astrophysics) luminosity functions for active galaxies.

These now agree (to a factor two) so the two selection techniques — hard X-ray emission and broad optical emission lines — do find the same population of objects.

The 'narrow emission line galaxies', once thought to be a different class, have the universal power law X-ray spectrum of slope 0.7 above 4 keV that all measured Seyfert 1 galaxies possess (Mushotzky). Also, better optical spectra showed faint broad wings on their Balmer lines, just like Seyfert 1 galaxies.

Some anomalies remain. No-one is sure whether or not Type 2 Seyfert 1 galaxies (differing from Seyfert 1 galaxies by the absence of broad hydrogen emission lines) have dragons in them. Also new Einstein identifications of active galaxy sources (Elvis, Harvard-Smithsonian Center for Astrophysics), include two that would not have been picked out by an optical survey since they have neither strong emission lines nor blue colours.

Much detailed theoretical work was presented, largely based on the concept of an accretion disk around a supermassive black hole. Although progress is being made in investigating these systems as demonstrated by, for example, Begelman (Berkeley), Duncan (Cornell) and Eilek (Institute of Mining and Technology, New Mexico), the models need to be explored further before the tenuous links between theory and observation can be strengthened. □

the presence of SV40 T antigen and that it occurs irrespective of whether the SV40 sequences are physically linked to the foreign DNA or merely co-injected with it.

The chief reasons for using viral vectors are to place foreign DNA sequences under the control of known viral regulatory elements, and to attach them to viral origins of replication so that the ectopic sequences can be amplified to high copy number in infected cells. Viral vectors have the further advantage that the foreign DNA can be delivered virtually synchronously to large populations of cells. Until recently, only one type of viral genome — that of SV40 — was used for this type of work. Typically, the late region of the viral DNA is replaced in whole or in part by foreign DNA sequences. Unlike bacteriophage λ vectors where the recombinant DNA molecules are packed *in vitro* into virus particles, the SV40-chimaeric DNAs are first mixed with a helper viral DNA and the mixture then used to infect permissive cells. Although only a small proportion of the cells can be infected with DNA, the susceptible cells efficiently take in both the recombinant DNA and its helper. The recombinant genome replicates and as long as it is not too large, becomes packed into viral particles. The resulting mixed virus stock can be used to infect enough cells to allow analysis of the expression of the foreign DNA by conventional biochemical techniques. Quite a large number of genes has now been propagated and expressed in this way. The most recent additions to the list are human growth hormone, hepatitis surface antigen, human genomic globin sequences (Pavakis, NIH) and the putative transforming protein, p21, coded by Harvey sarcoma virus (Khoury, NIH). In many cases the expression of the exogenous genes is surprisingly efficient, even when the foreign DNA is inserted into unpromising positions in the SV40 genome. For example, when a form of the rat preproinsulin gene from which all splice signals have been deleted is inserted into an SV40 vector which also lacks splice signals, recombinant mRNA is synthesized in large quantities under the control of the SV40 late promoter and translated efficiently (Gruss, NIH, see *Proc natn Acad Sci USA* 133; 1981). This result indicates that lack of splicing may not be as strong a deterrent as had previously been thought to the expression of ectopic DNAs.

Promising though the SV40 system is, it has suffered from two disadvantages. First, virus populations which contain recombinant genomes are always mixed, and sometimes the particles which carry foreign DNA sequences form only a small proportion of the virus stock. Second, the size of the foreign insertion is limited by the amount of DNA which can be packed into SV40 particles. However, several new vectors are being developed which promise to alleviate these difficulties. Bovine papilloma virus is capable of transforming

Designer genes

from a correspondent

SMALL MEETINGS on molecular aspects of biology are rare luxuries these days. Almost as fast as they well up, new ideas and discoveries become part of the mainstream, where they sustain shoals of workers most of whom are darting in the same direction at the same time. Consequently, it was refreshing to attend a small meeting* on 'Mammalian Vectors' during which the forty or so participants discussed work in varying stages of accomplishment which dealt, in surprising variety, with the problems of obtaining expression of ectopic genes in mammalian cells.

There are three methods in common use by which foreign genes (generally in the form of cloned fragments of genomic DNA or cDNA copies of mRNA) can be reintroduced into cultured mammalian cells: by microinjection, linked to segments of viral DNA, and by simultaneous transformation with two separate pieces of DNA — one carrying the gene of interest, the other a genetic marker which confers a

selective advantage on the recipient cells.

Each method has its own set of advantages and disadvantages and the three techniques therefore should be regarded as complementary rather than exclusive. Microinjection is an efficient way of introducing known quantities of DNA into cells whose exact location on the culture dish is known. Enough cells can be injected to allow detection by immunofluorescence of the products of expression of the injected DNA. Normally such expression is transient, lasting, at the most, for a few days. However, about one to two per cent of the injected cells can be developed into lines which express the foreign genes constitutively (Capecchi, University of Utah, Lo, University of Pennsylvania, French, Anderson, NIH). Furthermore, Capecchi showed that the efficiency with which constitutive cell lines are established can be increased some 20-fold, by injecting together with the foreign gene, a segment of DNA which spans the origin of simian virus 40 (SV40) DNA replication. The cause of this improvement is unknown, it is especially puzzling that the effect does not depend on

*Held from December 11 to 13 at the Banbury Conference Center of Cold Spring Harbor Laboratory.

mouse cells to yield cell lines which, by contrast to similar cells transformed by SV40, contain 10–120 copies of circular viral DNA with no evidence of integrated viral sequences. Howley (NIH) has attached the rat preproinsulin gene to the transforming region of bovine papilloma virus DNA and has established transformed cell lines which contain episomal copies of the recombinant DNA. The cells contain spliced preproinsulin mRNA and secrete significant quantities of proinsulin.

Gluzman (Cold Spring Harbor Laboratory) has developed lines of permissive simian cells which constitutively express functional SV40 T antigen. They can therefore be used in the absence of helper to propagate recombinant SV40 molecules in which the early region of the viral genome has been replaced by foreign DNA sequences, or to amplify DNA molecules which carry the SV40 origin. Using this system, Mellon (California Institute of Technology) has attached to the SV40 origin cloned copies of human chromosomal globin genes. Forty-eight hours after the introduction of these chimaeric molecules to COS cells, the globin genes direct the synthesis of approximately 10^4 molecules of globin mRNA per cell. The mRNAs are spliced, polyadenylated and have 3' and 5' ends identical to those of normal globin RNA. These results show that cloned copies of chromosomal genes which carry their own control sequences and which normally are active only in specialized cells can be efficiently expressed when reintroduced to undifferentiated tissue culture cells and amplified to high copy number as part of an SV40 recombinant genome.

Two groups (Thummell, University of California, Berkeley; Solnick, Cold Spring Harbor Laboratory) have used defective adenovirus 2 genomes to propagate and express foreign DNA. Using different techniques, both groups have constructed a series of recombinants in which the expression of the early region of SV40 is placed under the control of the major adenovirus late promoter. Many of the recombinants synthesize large quantities of authentic SV40 large T antigen. This result opens the way to use of adenovirus as a vehicle for the propagation and expression of moderate-sized eukaryotic genes, as, at least in theory, adenoviral vectors can be constructed which are able to accept segments of foreign DNA 30–35 kilobases in length.

Finally, although other selective markers are being developed (methotrexate resistance, neomycin resistance, dependence on guanine phosphotransferase; Southern and Subramanian, Stanford University), thymidine kinase, cloned either from herpes simplex virus or from chicken DNA, remains the marker of choice in co-transformation experiments. Using this technique, there is no difficulty in obtaining cell lines which contain a few copies of the foreign DNA of interest, in

the form of a large catenate formed by recombination between the incoming thymidine kinase gene, the carrier DNA and the foreign sequences themselves (Wigler, Cold Spring Harbor Laboratory, Ruddle, Yale). In general the level of expression of the foreign sequences is too low to be useful. Nevertheless, the system can be manipulated in interesting ways. For example, Kurtz (Cold Spring Harbor Laboratory) reported that the α -2 μ globulin genes which code for the major

urinary protein of rats can be introduced by co-transformation into cultured mouse cells. Several cell lines were obtained in which these foreign genes could be turned on by dexamethazone. The finding that co-transformed genes can continue to respond to hormones opens up the exciting possibility of analysing hormonally controlled genes by the same techniques that have been so successful in dissecting the fine structure of other viral and cellular promoters. □

Astrophysics and particle physics come together

from Frank E. Close

THE advances made in high-energy physics and astrophysics have enabled the 1970s to be described as the greatest ever in fundamental research. The discoveries of the J/ψ particle in 1974 and of charm in 1976 provided major evidence supporting the theory that united the electromagnetic and weak forces. Some of these particles' properties caused attention to be focused on a candidate theory of quark interaction, named 'quantum chromodynamics' or simply QCD. (The so-called 'strong' nuclear force is now viewed as a complicated manifestation of this more fundamental QCD force which acts on the quarks — the constituents of the subnuclear particles.) Although still some way from being verified, accumulating data do appear to fit in with theoretical expectations based on QCD.

The mathematical similarities among the theories of the electromagnetic, weak and 'strong' forces are so striking that theorists have tried to build a grand unifying theory which would include all the natural forces except gravity. The simplest way of implementing the theory has already successfully predicted the existence and mass of the fifth variety of quark (the so-called 'bottom' quark). Much experimental effort is now centred on the prediction that protons are unstable.

These topics were all reviewed at the annual meeting of the UK High Energy Theoretical Physicists held in December at the Rutherford and Appleton Laboratories.

J. Badier (Paris) reviewed experimental data on the behaviour of quarks inside protons and pions and together with R. Roberts (Rutherford Laboratory and CERN) discussed how well these and other data fit with the QCD theory of quark interactions. The answers continue to be very encouraging. Everything seems to fit with the theory, and although far from proved one increasingly feels that it would. *F. E. Close is at the Rutherford and Appleton Laboratories, Didcot, UK.*

now require something totally unexpected to overthrow it.

Although the basic tenets of QCD are rather simple, the rich structure arising from them is still only poorly understood. M. Creutz (Brookhaven, New York) described some computer-based studies of QCD performed on a 'lattice' (where space-time is treated as discrete rather than as a continuum) which suggest that QCD leads to confinement of gluons (the QCD quanta, analogues of photons in QED) and thus, it is speculated, of quarks. Direct proof of this is still, however, unavailable. Some technical problems on the behaviour of QCD in the IR region were discussed by J. C. Taylor (Cambridge). Some feel that deeper understanding of this domain might resolve the confinement question.

Looking to future developments, searches for proton decays and their significance for grand unified theories of the natural forces were discussed. D. Perkins (Oxford) described the experiments planned or in operation. Present limits show that the proton lifetime is greater than some 10^{29} years. Perkins pointed out that logistic problems with experiments deep underground and with backgrounds from neutrinos would probably prevent this limit ever being pushed beyond 10^{33} years. Grand unified theories predict 10^{31+2} years — satisfyingly in this window! Observation of such proton decays in the near future could provide a major impetus for belief in grand unified theories.

The theoretical issues involved in constructing grand unified theories of the particles and forces were detailed by J. Ellis (CERN). In our cold world, energy densities are rarely much greater than 1 GeV per fm³ and the forces appear quite different from one another in their strengths and properties (see Dombey *Nature* 288, 643, 1980). The unification of the forces is manifested at energy densities some 10^{15} times greater than these (note

that gravity is not important until densities of yet a further 10^4 are met, and so gravity can still be ignored) High-energy accelerators of subatomic particles can create locally energetic conditions and do appear to show a glimpse of the trends required if QCD is to be unified with the electroweak force, but we are still many orders of magnitude from 10^{15} GeV. However, such energy densities were probably common in the early Universe and it is here that a new merging of interests in high energy and astrophysics is taking place. These were described by Ellis and R. Taylor (Sussex)

Proton stability is particularly significant as several years ago Andrei Sakharov (Gorki) had noted that three ingredients were required if we were to understand the apparent excess of matter over antimatter in the Universe at large. One is an absolute distinction between them: this is indeed provided in nature by the phenomenon of CP violation (whose discovery led to the 1980 Nobel prize for Cronin and Fitch of the US). The second is an absence of thermal equilibrium provided in modern cosmology theory by the big bang expansion. Finally, proton instability is required. This is indeed provided by the grand unified theories; we wait to see if it is also provided by nature!

Within the grand unified theories it has been argued that mechanisms naturally exist that will give rise to the observed quantitative matter excess. Ellis pointed out that the validity of this 'explanation' can be tested because the mechanism also generates an electric dipole moment for the neutron. At present there is an experimental upper limit that places this moment at less than about 10^{-24} e cm. Present experiments hope to lower this by three orders of magnitude. Ellis noted that the above theories provide a 'cosmological lower bound' some four orders of magnitude below the present limit! This should provide a significant spur to experiment and provides another example of how studies of subnuclear particles can reveal the mechanisms responsible for the development of the early Universe.

One of the more recent lines of investigation is into the question of whether there are new phenomena between the present energies and 10^{15} GeV (apart from the predicted production of massive W, Z particles of the electroweak theory and the postulated sixth quark and Higgs particles). In particular, the nature of the Higgs particles is unclear.

During the past year there has been much interest in the possibility that the Higgs particles are not elementary but are dynamically bound composite systems of more fundamental entities. E. Eichten (Harvard) described so-called 'hypercolour' or 'technicolour' theories that contain the Higgs particles as composite bound states and in some formulations require substructure to the leptons and quarks. These theories predict that a new ultra-

high-energy scale of phenomena exists in nature. Eichten described possibilities of testing these ideas at present energies. Experimentalists are urged to search for the light pseudo-Goldstone bosons predicted in the theory (possibly of a mass of about 10–20 GeV and producible in electron-positron annihilations) or to search for the rare decay modes that technicolour's existence can induce from the conventional particles (like $\mu \rightarrow e\gamma$, $K \rightarrow \mu e$).

Finally, in all these brave schemes to unify our understanding we come up against the perennial problem — gravity. How gravity fits into nature's scheme is still

something of a mystery. P. van Nieuwenhuizen (Stony Brook, US), one of the founders of 'supergravity' theory, described developments in this exciting field. Some of the technical problems that arise attempts to build a satisfactory quantum theory of gravity appear to be overcome if gravity is but part of a larger theory — supergravity. This theory is very elegant mathematically and yet, frustratingly, we have little or no indication that Nature has made use of it. van Nieuwenhuizen succinctly summarized the beauty of the mathematical structures by saying "I hope that Nature is aware of our efforts" □

Υ physics—from bottom to top

from Roger Cashmore

THE Υ meson was first observed in proton-proton collisions late in 1977 at Fermi National Accelerator Laboratory (FNAL). The meson had a mass of ~ 9.45 GeV and was seen in its decays to $\mu^+\mu^-$ and e^+e^- with a width consistent with the measurement resolution. (The width of a resonant state is related to its lifetime. Very short lifetimes lead to large widths. However, even a narrow state can appear to be broad if the experiment is limited by the measurement accuracy. The width then reflects the experimental resolution.) A further state, the Υ' , was also indicated at a mass of ~ 10.00 GeV.

These observations were quickly confirmed in experiments at the e^+e^- storage ring, DORIS, at DESY in Germany. The Υ and Υ' were observed as separate states, ~ 580 MeV different in mass, and since then the Υ has been shown to have a total width of 20–50 keV, a remarkable fact for such a high-mass state and rather reminiscent of the ψ . Generally, states of such high mass decay quickly into many lower-mass particles leading to large intrinsic widths of hundreds of MeV. Combined with the FNAL data, these accurate measurements of the masses of the Υ and Υ' states suggested the existence of yet a third Υ resonance at a mass just beyond the reach of the DORIS storage ring.

It appeared that the bound state system of a new quark, the charge $\frac{1}{3}$ bottom quark (b), and its antiparticle ($\bar{b}$) was being observed. The Υ was the ground-state $b\bar{b}$ meson and the Υ' its first excited state. New quarks, bottom (charge $\frac{1}{3}$) and top (charge $\frac{2}{3}$), had been implied by the earlier observation of a new lepton, the τ lepton, at the electron-positron storage rings SPEAR and DORIS, capable of reaching 7 and 10 GeV energies respectively.

The past year has seen the successful operation of a new electron-positron storage ring, CESR, at Cornell which can be used for experiments in the 8–16 GeV mass region. This covers the Υ mass range and allows much more detailed studies of this new quark system. The Υ

and Υ' states were confirmed and the third state, the Υ'' , observed at a mass of ~ 10.32 GeV. All three states had widths consistent with the machine resolution, suggesting that they were all narrow resonances.

However, as the beam energies were raised a fourth structure, the Υ''' , was observed at a mass of ~ 10.54 GeV but this had a width of ~ 10 MeV, larger than the machine resolution and in marked contrast to the other states. This suggests that the state is above the threshold for decay into $B\bar{B}$, B being a $b\bar{q}$ meson (where q is one of the more familiar u, d, s quarks) carrying the flavour of the new b quark. Thus these Υ states provide a laboratory for the study of both the bound systems of this quark, $b\bar{b}$, and the open flavour states $b\bar{q}$.

These two areas are vital in validating the present descriptions of the forces in particle physics. The strong force, which binds the quarks and antiquarks to form the mesons, is hypothesized to be indifferent to the quark flavour and only related to their colour through the exchange of coloured gluons between them. (Earlier results had indicated the need for three colours of each flavour of quark.) Hence $q\bar{q}$ systems should be similarly independent of the quark flavour. This model for the strong interaction is known as QCD (quantum chromodynamics). On the other hand, the weak force, now believed to be unified with the electromagnetic interaction according to the gauge theory of Weinberg and Salam, implies specific transitions between the different flavours of quark, for example, $c \rightarrow s$, $s \rightarrow u$. These hypotheses need quantitative verification.

In the $b\bar{b}$ bound system the mass difference of the Υ and Υ' is close to that of the ψ and ψ' , indicating that, although the new quark has a mass ~ 5 GeV in contrast to the charm quark mass of ~ 1.5 GeV, the quark forces (or potentials) are essentially

Roger Cashmore is in the Department of Nuclear Physics, University of Oxford

unaltered. This is the remarkable result, that the force is flavour independent, which is expected if the strong forces are indeed due to the exchange of coloured gluons as embodied in QCD.

These narrow states, too low in mass to decay to $b\bar{q} + \bar{b}q$ mesons, are further expected in QCD to decay by the emission of three gluons (the smallest number allowed by a variety of selection rules). Detailed analysis of the hadronic final state of the T indicates that even this prediction is consistent with all of the observations. Thus this T system confirms dramatically the QCD model of strong interactions.

The T'' should be a prolific source of B mesons, and hence detailed studies of their properties (and those of the b quark) will be possible. A B meson has yet to be demonstrated at Cornell (an earlier possibility from a CERN experiment has since disappeared with further data). Its decay modes are many and the multiplicity of particles in each is expected to be very high, making a clean identification difficult. Nonetheless, it is possible to look for irregularities that might be associated with the T'' and hence the B particles. An excess of K mesons is observed and these are expected if the b quark (within the B meson) undergoes a weak decay transition to the charm quark, which has a well documented predilection for strange particles in its decay. The excess is, within the large errors, just what would be expected. Furthermore, the semi-leptonic weak decays of the B mesons ($B \rightarrow \nu e \nu$, where ν represents one or more hadrons) lead to increased numbers of electrons and muons. Such increases are observed and the inferred branching ratios for these decays are 24 ± 6 per cent, whereas the standard model, containing a top quark, would suggest a value of ~ 34 per cent, not widely inconsistent with the experimental value. Thus the data, sparse as they are at the moment, suggest that the standard model remains intact with the most likely existence of a top (charge $\frac{2}{3}$) quark and a weak decay chain $t \rightarrow b \rightarrow c \rightarrow s$. Higher energies than those achieved so far at PETRA (~ 38.6 GeV) will be needed in the search for this quark. A very recent paper (S. Glashow *Phys. Rev. Lett.* **45**, 1914, 1980) puts forward further theoretical arguments for a mass of 38 ± 2 GeV for the top quark. Experimentalists should soon be able to check on this. The present gauge model relies on its existence. The new facilities in Cornell will undoubtedly produce far more quantitative results in the future. These will allow more systematic tests of our current ideas of the quarks and their strong and weak interactions and perhaps in this quantitative exposure new possibilities will emerge. After all, we have as yet no explanation for the number of generations of quarks and leptons. Experiments will perhaps one day provide the clue. An exciting few years are in store at Cornell. $\square$

Molecular hydrogen in T Tauri

from Stephen P. Maran

THE discovery of ultraviolet molecular hydrogen emission in T Tauri with the International Ultraviolet Explorer satellite (see Brown *et al.* this issue of *Nature* p 34) has considerable potential for three fields of research: the nature of T Tauri stars and their associated nebulae, the formation and development of the Solar System, and the study of molecular hydrogen in the Galaxy. T Tauri stars have masses about equal to that of the Sun, but are rotating much more rapidly and are thought to be stars still-in-the-making, only about a million years old and still contracting towards the Main Sequence. Some appear to be losing mass, the so-called 'T Tauri wind', while others appear to be still accreting matter from the surrounding nebulae. It is widely accepted that the Sun passed through a T Tauri phase and models of Solar System formation take into account the postulated high wind and luminosity and the concomitant decreasing gravitational attraction of the early Sun to solve a number of problems. Among them are the placing of Mercury and perhaps other planets into their present orbits, the sweeping out of excess gas and dust in the protoplanetary nebula, the dispersal (in modern versions of the Laplace nebular hypothesis) of gas rings in the planetary orbits, the formation of chondrules found in meteorites, and the reduction of the Sun's rotation to its present very low rate. T Tau itself, the prototype of the stellar class, is associated with a small luminous cloud called Burnham's nebula.

It had been thought unlikely that ultraviolet emission known to occur in nebulae would be detectable in Burnham's nebula with instruments having the sensitivities of the two IUE spectrographs. Brown *et al.* confirmed this, for they found no previously known nebular lines in the spectrum of the object, but instead discovered emission in the H_2 Lyman bands that previously were seen only in sunspots. In recent years, astronomers have concluded that the interstellar gas is not dominated simply by atomic hydrogen but that there is also a substantial molecular component. This has been largely inferred from observations of the CO molecule in the galactic plane but H_2 in the general medium has also been seen through its ultraviolet absorption lines (a technique feasible only in the directions of appropriate hot stars that serve as background continuum sources), and H_2 in several nebulae (including Burnham's nebula) has been found via the infrared

vibration bands. More recently, a bound-free absorption continuum of the H_2^+ ion has been detected in planetary nebulae (S. Heap and T. P. Stecher *Proc. Conf. on the Universe at Ultraviolet Wavelengths* 1980).

Brown *et al.* distinguish between two mechanisms for H_2 excitation in Burnham's nebula and favour the 'shock excitation' model. The nebula is much smaller than the Oort cometary cloud that surrounds the Sun, and hence phenomena on its scale may well be relevant to the early evolution of the Solar System. Since H_2 must be a principal component of the nebula, observations with more sensitive ultraviolet instruments having very fine angular resolution (notably, the High Resolution Spectrograph now under development for flight on the *Space Telescope*) will enable astronomers to investigate the dynamics of the shocked gas. In the meantime, nebular specialists will be busy calculating where else in the Galaxy this new diagnostic may be detectable. $\square$

Planarian neoblasts

from J. Baguña

IN a News and Views article on 21 August 1980 Jonathan Slack¹ reviewed briefly the sources of the cells involved in regeneration in planaria and, in doing so, made several assertions that deserve comment.

(1) 'Neoblasts' are indeed small basophilic cells present in planaria. But to think of them as cells 'reserved' for regeneration is equivocal since many planarian species, having sexual reproduction, never regenerate in nature. So what are the neoblasts for? The answer lies in considering the neoblast not just as a 'reserve' cell but as the stem cell of most differentiated cell types supplying transit cells to balance the death of cells.

(2) In planarian regeneration it is meaningless to argue whether neoblasts come from all parts of the body or from local tissues through metaplasia because the blastema can be formed (and is formed) from pre-existing local neoblasts proliferating near the wound. Moreover, Wolff's experiments², which have been cited as evidence for migratory neoblasts, have been consistently criticized³⁻⁶ because at most, they only show the 'repopulation' of the irradiated region by healthy neoblasts through cell division and do not prove a direct migration of neoblasts from unirradiated regions to the wound.

(3) Slack points out that "histological and electron microscopic studies suggest that local dedifferentiation does occur." The

Stephen P. Maran is Senior Staff Scientist in the Laboratory for Astronomy and Solar Physics, NASA-Goddard Space Flight Center, Maryland.

J. Baguña is in the Department of Genetics, University of Barcelona, Spain.

available evidence is scanty and unclear so this seems a bold assertion. Other authors^{7,8} have been unable to find any dedifferentiating features in any of the blastema constituting cells.

(4) The results reported by Gremigni and Miceli⁹ supporting cell dedifferentiation and transdifferentiation in planarians, though offering new ways to re-evaluate the problem of blastema formation, are not as conclusive as has been suggested. First of all, the sequence proposed: germ cell → blastema cell → muscle cell, and its corollary of dedifferentiation and metaplasia are not necessarily related as the loss of an haploid complement during spermiogenesis, though one of the first steps from neoblasts to male germ cells, is only a trivial phenomenon regarding cell determination. Unless spermatids or primary spermatocytes are shown to give rise to blastema cells, and then to differentiated cells, we cannot talk properly of dedifferentiation and metaplasia. At the most, this would qualify for dedetermination (or trans-determination) and the sequence proposed should be read instead as predetermined or determined: germ cell → blastema cell → muscle cell. Second, and more important, the sequence suggested by Gremigni and Miceli, if proved correct, has to be demonstrated for other differentiated cell types since several planarian species, being asexual, lack germ cells. Besides, it is possible, as suggested by Betchaku⁴ and Pedersen⁸, that both processes (neoblasts and metaplasia) occur during planarian regeneration. If this were the case, the important issue would be to know the partitioning between both phenomena during planarian regeneration.

(5) Though most (but not all, witness *Hydra*) regenerative phenomena in the Animal Kingdom have been proved to occur through cell dedifferentiation and re-differentiation, the present trend to extrapolate this mechanism to all animal groups (including planarians) overlooks the possible relationship between the actual mechanism of regeneration and the complexity of the species. Planarians are one of the most primitive acoelomates, and their tissular complexity is rather low if compared with other epimorphic-regenerating organisms like Amphibia or Insecta. No wonder that dissimilar mechanisms may operate during regeneration in these groups. The planarian neoblast system, based on a unique and multipotential self-renewing stem cell, though appropriate to the low-level complexity of planarians, is clearly

inadequate for organisms with static (decaying) cells and transit cells that require different determined stem cells placed in different body regions. Consequently, the dedifferentiation process, necessary for Amphibia regeneration, is not necessary for organisms like planarians which are in a continuous state of rapid cell turnover¹⁰. Unless new data on planarian histology and cytology, cell dynamics, and life cycle are obtained, the extrapolation of other mechanisms of regeneration to planarians^{1,9} and the construction of theoretical models of planarian regeneration¹¹ will be of little value.

J.M.W. Slack comments: In my article I attempted to show that several *independent* issues are involved in establishing the source of cells for regeneration, each of which must be approached through a different experiment. As Baguña implies, the simple dichotomy — 'neoblasts versus metaplasia' — is often not an adequate statement of the problem. I believe that the controversy which has arisen in this field in the past has largely been caused by an imprecision in the question asked. □

J.M.W. Slack is in the Imperial Cancer Research Fund Laboratories, Mill Hill, London

A carbohydrate signal for intracellular transit

from Mike Geisow

JUST over a year ago, the function of the oligosaccharide chains added to many newly synthesized proteins was discussed in these columns (*Sugars and intracellular recognition*, *Nature* 281, 15, 1979). It was noted that an early hypothesis — that carbohydrate acted as a marker for secretion — had been superseded by the discovery of the signal peptide. However, it did seem as if carbohydrate on non-secreted proteins might still act as a signal for intracellular transit from the endoplasmic reticulum to subcellular, membrane-enclosed destinations. Recent work has shown that this is indeed the case for lysosomal enzymes.

Characteristically, lysosomal enzymes are glycoproteins and many have been shown to contain mannose residues phosphorylated at carbon-6 in the hexose rings. Both glycosylation and phosphorylation occur in the endoplasmic reticulum and appear to be directed by the sequence and conformation of the polypeptide substrate, not all proteins receive both modifications. The importance of both the

mannose and the phosphate groups in receptor-mediated processes first became apparent when glycoproteins expressing both determinants were shown to be rapidly and specifically endocytosed by fibroblasts. Next, the uptake of a variety of lysosomal hydrolases was found to be subject to competitive inhibition by mannose phosphate. These observations suggested three rival routes for lysosomal enzymes from their sites of synthesis to primary lysosomes. Two of these, involving secretion followed by re-uptake or cycling as receptor complexes via the surface membrane, appear to be of minor importance: receptor-saturating levels of extracellular mannose-6-phosphate do not deplete lysosomal enzyme levels. Fischer and his colleagues (*J. biol. Chem.* 255, 9608, 1980) now present firm evidence for a major intracellular transit system using β -hexosaminidase B in fibroblasts.

It was found that 80 per cent of the mannose-6-phosphate receptors were intracellular and faced the interior of endoplasmic reticulum or lysosomes. The highest specific activity of enzyme-binding receptors was present in these cell fractions. Moreover, the intracellular receptors were combined with endogenous enzymes and presented an occupancy gradient which ran steeply downhill from endoplasmic reticulum to lysosomes. The inference is that proteins which receive phosphorylmannose in the endoplasmic reticulum as a result of as yet unknown conformational signals in the polypeptide, are recognized and transferred to primary lysosomes by an essentially intracellular route.

In the report referred to earlier, attention was drawn to the frequent observation of secretory protein proteolysis in endocrine cells treated with inhibitors of protein glycosylation. In such cells, failure of the glycosylation system presumably produces not only non-glycosylated hormones, but also carbohydrate-free acid hydrolases co-synthesized in the endoplasmic reticulum. What happens to the non-glycosylated lysosomal enzymes? In I-cell disease, defective addition of phosphorylated carbohydrate to the enzymes results in their secretion. Co-packaging of acid hydrolases and normal secretory products in the acid environment of a secretion granule may well explain the observed aberrant proteolysis. This argument leads to an interesting thought — is such proteolytic attack on hormones always abnormal? Some hydrolases appear to be naturally present in secretion granules as opposed to lysosomes. These are the conversion enzymes which give rise to products as diverse as insulin, oxytocin and enkephalins from the original precursors. Do such proteases arrive at their destination in the secretory granule because they normally lack a mannose-6-phosphate-containing oligosaccharide chain? □

Mike Geisow is in the National Institute for Medical Research, Mill Hill, London

- Slack, J. M. W. *Nature* 286, 760 (1980).
- Wolff, E. in *Regeneration* (ed. Rudnick, D.) 53 (Ronald Press Co., New York 1962).
- Lindh, N. O. *Ark. Zool.* 10, 497 (1957).
- Betchaku, T. *J. exp. Zool.* 164, 407 (1967).
- Sugino, P. H., Okuno, Y. & Yoshinobu, I. *Mem. Osaka & Kyoto Univ.* 19, 63 (1976).
- Hori, I. *J. Electron Microsc.* 27, 89 (1978).
- Pedersen, K. J. *Wilhelm Roux's Arch.* 169, 134 (1972).
- Gremigni, V. & Miceli, C. *Wilhelm Roux's Arch.* 188, 107 (1980).
- Baguña, J. *J. exp. Zool.* 190, 177 (1974).
- Slack, J. M. W. *J. theor. Biol.* 82, 105 (1980).

Beyond the giga-seal

from Robert McBurney

TWO recent technical developments promise a new era in the examination of the movements of ions across the surface membranes of excitable cells. As well as the achievement of much higher current resolutions it is expected that one technique, using an 'excised patch', will make it possible to monitor changes in membrane properties while being able to directly influence the environment on the inside surface of the cell membrane.

The new techniques stem from the development by Neher and Sakmann (*Nature* **286**, 71, 1980) of a method for the direct measurement of the ionic current flowing through single ion channels in biological membranes. These recordings of picoampere (10^{-12} A) currents are made with a specially constructed glass pipette which seals off a small patch of membrane (about $1\ \mu\text{m}$ in diameter) on the surface of a cell from the rest of the surface membrane. The successful interaction of agonist molecules in the solution inside the pipette with receptor-ion channel complexes on the patch of membrane generates discrete pulses of ionic current which are recorded by a low-noise current-to-voltage converter. The current loop for measurements is from the pipette internal solution through open ion channels in the patch to the cytoplasm, then through 'leakage' channels in the membrane outside the patch to the bulk bathing solution and through the recording device to the pipette solution.

If some of the current flowing through the ion channels fails to complete this recording circuit, and a proportion returns via the seal between the pipette wall and membrane or through leakage channels in the patch itself, the true single-channel current amplitude is not measured. While the latter return pathway is not significant, except for larger patches on very small cells, the achievement of adequate seals has been a key factor in the application of this technique to different preparations. Not only does the relationship between seal resistance (R_{seal}) and recording system resistance determine the proportion of the actual current recorded but the thermal current noise in the seal resistance (inversely proportional to $\sqrt{R_{\text{seal}}}$) determines the lower limit of current resolution of the recording system.

Despite problems in creating adequate seals, the patch recording technique has been successfully applied over the past five years to a variety of preparations, including denervated frog muscle, cultured rat muscle (myoballs), locust muscle and

squid giant axons. The channels investigated have mostly been those activated by receptor agonists (cholinergic in frog and rat muscle, glutaminergic in locust muscle), although in one case (squid giant axon) the properties of voltage-dependent K^+ channels have been studied. The technique has been used to confirm and extend data obtained by the analysis of membrane current fluctuations (see Stevens *Nature* **270**, 391, 1977). In addition, it has yielded new information concerning the detailed kinetics of action of some drugs, for example local anaesthetics on acetylcholine-activated ion channels (Neher and Steinbach *J Physiol, Lond* **277**, 153, 1978), and has revealed the behaviour of ion channels during desensitization (Sakmann *et al Nature* **286**, 71, 1980).

The first exciting new development to come from Neher and Sakmann's technique arose from efforts to increase the sealing resistance. By applying suction to the recording pipette as it was placed on the surface membrane of a cultured muscle cell, F Sigworth and E Neher (Göttingen) found that a very high-resistance seal ($>10^9\ \Omega$) was, in some cases, suddenly created. This seal and its associated low current-noise even at a bandwidth from DC-7 kHz allowed Sigworth and Neher to resolve single voltage-dependent Na^+ channels: those normally associated with the generation of muscle action potentials. In their preliminary report (*Nature* **187**, 447, 1980), as well as describing the 'giga-seal' and presenting clear records of single Na^+ channels, they have shown that these channels have conductance and average lifetime values which agree with those obtained by more indirect means. They also demonstrate the independence of channel behaviour by showing that the sum of many records, usually containing just one channel opening in response to a fixed membrane depolarization, mimics the behaviour of many channels operating simultaneously.

The future application of the 'giga-seal' technique to excitable tissues will permit the investigation of 'gated' ion channels in cells or regions of cells not amenable to voltage-clamp analysis with intracellular microelectrodes, such as smooth cells, gland cells, unmyelinated axons and possibly even nerve terminals. In addition, the technique provides improved current resolution and, for single-channel currents already observed by the original method, a much wider bandwidth for recording which will undoubtedly reveal finer details in the kinetics of channel activation and inactivation.

Perhaps more exciting than the achievement of the 'giga-seal' itself has

been the recent observation made by two groups of workers, R Horn and J Patlak (Yale) and O Hamill and B Sakmann (Göttingen). They found that, once established, this seal was almost impossible to shake off. When the electrode was pulled away from the cell it seemed that the patch of membrane was excised on, or within, the tip of the electrode and single-channel currents could still be observed if the transmembrane voltage was appropriately adjusted. After a short time the recording ceased, possibly due to a membrane-bound vesicle being formed within the electrode tip (because the high-resistance seal remained). The two groups solved the problem of vesicle formation in different ways. Horn and Patlak (*Proc natn Acad Sci USA* **77**, 6930, 1980) found that fluoride ions delayed or prevented the process, while Hamill and Sakmann (*J Physiol, Lond* in the press) found that by lifting the electrode through an air-water interface they were able to re-expose the patch.

In their preliminary report Horn and Patlak illustrate one of the special features of this preparation: they examine the dependence of the conductance of acetylcholine-activated channels on sodium ion concentration in a situation where the latter could be defined on both sides of the membrane. Such access permits a vast array of experiments to be undertaken with controlled environments on either side of the membrane, similar in principle to the study of ion transport in 'doped' artificial lipid membranes, and these will certainly result in a clearer idea of the mechanism of ion transport and selectivity of natural ionophores.

However, the most important aspect of the preparation may simply prove to be access to the inside surface of cell membranes. It will now be possible to investigate directly the action of intracellular messengers (for example, cyclic nucleotides, Ca^{2+}) and cellular metabolic products on membrane excitability and transport processes. It is hoped that this 'excised patch' technique will be generally applicable to cells derived from a wide range of tissues. For example, its application to heart cells will certainly provide an excellent preparation with which to study the relationships between 'second messengers', and protein phosphorylation and ion channels in the hormonal modulation of membrane conductance changes underlying cardiac pacemaker activity. □

Erratum In *Antimatter back to front* by C H Llewellyn Smith (*Nature* **289**, 534 February 12, 1981) the last sentence in the first paragraph should have read "The discovery of CP violation by Christenson, Cronin, Fitch and Turlay in 1964, for which Cronin and Fitch were awarded the 1980 Nobel prize, was therefore almost as great a shock as the discovery of parity violation".

Corrigendum. In *Cement — a respectable material?* by D D Double (**289**, 348, January 29, 1981) the figure of "100 MN m⁻²" quoted as the modulus of mother-of-pearl shell should have been "about 60 GN m⁻²".

Robert McBurney is in the Department of Physiology, The Medical School, Newcastle upon Tyne

ARTICLES

Mesoscale convection in the clouds of Venus

C. C. Covey & G. Schubert

Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

The cellular structures seen in UV images of the atmosphere of Venus probably represent convection cells. If so, this shape (horizontal dimensions 10–100 times greater than vertical dimensions) cannot be accounted for by radiation transfer alone but requires anisotropic eddy diffusivity

UV PICTURES of Venus taken by the Mariner 10 and Pioneer Venus spacecraft show cellular structures near and downwind of the subsolar point at cloud top heights of 60–70 km (Fig. 1)^{1–3}. These structures may result from various eddy phenomena, for example, wave trains intersecting to form a criss-cross pattern. However, the most natural interpretation, given their proximity to the subsolar point, is that they are atmospheric convection cells visible at the cloud tops.

Cloud level convection

The presence or absence of an adiabatic layer at the cloud tops would be strong evidence either for or against convection in this region of the atmosphere. Unfortunately, the four Pioneer Venus probes entered the atmosphere away from the subsolar point, at night or in the early morning. They, nevertheless, all measured an adiabatic temperature gradient between ~50 and 55 km altitude with stable layers above and below⁴. We therefore regard the UV cellular features as evidence for convection in the cloud-top subsolar region and suggest that the cloud level adiabatic layer in this region is either higher or thicker than the one measured by the probes, so as to encompass the cloud tops⁵. For example, in the 55–60 km altitude range the static stability measured by the probes has a mean value of 2–3 K km⁻¹. This 5-km thick layer of the atmosphere could become adiabatic if the temperature were increased by 10–15 K between morning and afternoon. Then the adiabatic layer would extend to 60 km, the lower limit for the altitude of the UV features. The required heating is compatible with a diurnal temperature cycle involving an atmospheric superrotation rate of 5 days⁶ and a radiative cooling time of ~100 days⁷, which represents a temperature variation of ~5%. Alternatively the cellular features may be located in an adiabatic layer above 65 km altitude (where static stability data do not exist)⁵. We might also be seeing penetration of lower lying convection into an overlying stable cloud top region².

An adiabatic cloud top layer is likely to be bounded above and below by regions of high static stability caused by radiative cooling of the cloud tops and radiative heating of the cloud base⁸. The depth of the convecting layer is then of the order of 10 km. Given the horizontal scale of the cellular UV features, 10²–10³ km, this implies aspect ratios (diameter/depth) of the order of 10–10². This is several times the aspect ratio of mesoscale cellular atmospheric convection on Earth⁹, which in turn is several times the aspect ratio obtained in laboratory and theoretical studies of cellular convection at moderately supercritical Rayleigh numbers¹⁰.

We now present a theory to explain these high aspect ratios for cloud level convection in the Venus atmosphere. Mesoscale atmospheric convection on Earth takes place immediately above the planetary surface and it is, therefore, affected by the

very steep temperature gradient in the thermal boundary layer¹¹. The properties of the cloud level Venus atmosphere together with the large aspect ratio of the cells require that the thermal diffusivity due to small-scale eddies be large and anisotropic.

Radiative and eddy processes

The major difference between laboratory and atmospheric convection is that in the latter radiative and eddy processes play a major part in the transport of heat and momentum. To represent the role of eddies an eddy viscosity ν and an eddy thermal diffusivity κ are used. These eddy diffusivities may be anisotropic. To account for this, we let ν and κ represent the diffusivities in the vertical direction and $m\nu$ and $m\kappa$ the horizontal diffusivities. Thus we assume that m , the ratio of horizontal to vertical diffusivity, is the same for momentum and heat transport. To account for radiative heat transfer within the convecting fluid, we linearize the Boussinesq equations for a compressible atmosphere with an internal heat source¹². Q about a basic state in radiative equilibrium. The incremental



Fig. 1 UV image of Venus from the Pioneer Venus Orbiter Cloud Photopolarimeter². The rotation of the atmosphere is from right to left. Cellular structures may be seen in the equatorial region, near and downwind of the subsolar point.

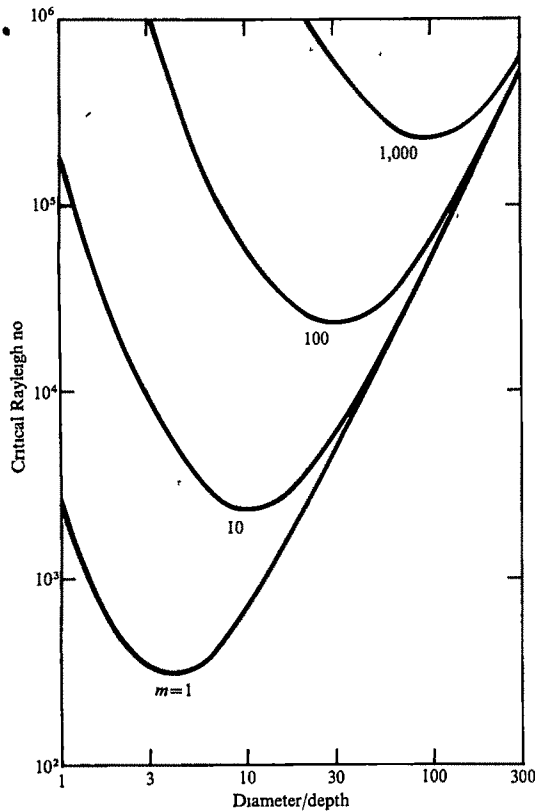


Fig. 2 Critical values of the Rayleigh number for the onset of convection as a function of aspect ratio, for different values of eddy anisotropy ratio m . For these curves $\tau = 10$ and $\gamma = 1$, as implied by $\kappa = 10^6 \text{ cm}^2 \text{ s}^{-1}$.

heating in the perturbed state Q' is approximated by the newtonian cooling law

$$Q'/\rho c \approx -\theta^*/\tau^* \quad (1)$$

where ρ is the density, c is the specific heat at constant pressure, θ^* is the perturbation temperature, and τ^* is the radiative cooling time. The newtonian cooling approximation is valid at cloud levels in the Venus atmosphere, and values of τ^* have been obtained⁷ at different altitudes from detailed radiative transfer calculations.

For simplicity we assume that the thermal expansion coefficient α , the basic state temperature gradient dT_0/dz^* (z^* is altitude), and the adiabatic temperature gradient Γ_{ad} are all constant. Then the governing equations for θ^* , perturbation velocity $\mathbf{u}^*$, and perturbation pressure π^* are

$$\left(\frac{\partial}{\partial t^*} - m\nu\Delta_2^* - \nu\frac{\partial^2}{\partial z^{*2}}\right)\mathbf{u}^* = -\nabla^*\pi^* + g\alpha\theta^*\mathbf{k} \quad (2)$$

$$\nabla^*\mathbf{u}^* = 0 \quad (3)$$

$$\left(\frac{\partial}{\partial t^*} + \frac{1}{\tau^*} - m\kappa\Delta_2^* - \kappa\frac{\partial^2}{\partial z^{*2}}\right)\theta^* = \left(\left|\frac{dT_0}{dz^*}\right| - \Gamma_{ad}\right)(\mathbf{k} \cdot \mathbf{u}^*) \quad (4)$$

In the above, $-g\mathbf{k}$ is the acceleration of gravity and Δ_2^* is the laplacian with respect to horizontal coordinates. We have assumed a motionless basic state and have thus neglected any wind velocity gradients. The zonal wind at the cloud tops is large but its vertical shear is unknown⁵. The fact that the putative convective features are cells instead of rolls argues that the wind shear is small enough to be neglected³.

We remove dimensions from equations (2)–(4) using the depth of the fluid d as the length scale, d^2/κ as the time scale, and $(|dT_0/dz^*| - \Gamma_{ad})d$ as the temperature scale. This gives

$$\left(\frac{1}{P}\frac{\partial}{\partial t} - m\Delta_2 - \frac{\partial^2}{\partial z^2}\right)\mathbf{u} = -\nabla\pi + R\theta\mathbf{k} \quad (5)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (6)$$

$$\left(\frac{\partial}{\partial t} + \frac{1}{\tau} - m\Delta_2 - \frac{\partial^2}{\partial z^2}\right)\theta = \mathbf{k} \cdot \mathbf{u} \quad (7)$$

where $\mathbf{u}$, π , θ , t , z , ∇ , and Δ_2 are dimensionless quantities and operators. Physical processes within the fluid are then determined by four nondimensional parameters: m , the Prandtl number

$$P \equiv \nu/\kappa \quad (8)$$

the Rayleigh number

$$R \equiv \frac{\alpha g(|dT_0/dz^*| - |\Gamma_{ad}|)d^4}{\nu\kappa} \quad (9)$$

and the dimensionless radiative cooling time

$$\tau = \tau^*\kappa/d^2 \quad (10)$$

We must also account for the effect of radiative transfer at the boundaries of the convecting fluid. At the upper boundary, the additional heat conducted upwards in the fluid $-\rho\kappa d\theta^*/dz$ must equal the extra heat radiated to space $4\sigma T_0^3\theta^*$. A similar condition applies at the lower boundary. Thus, in nondimensional form,

$$\frac{d\theta}{dz} \pm \gamma\theta = 0 \quad \text{at } z = \pm \frac{1}{2} \quad (11)$$

where

$$\gamma \equiv \frac{4\sigma T_0^3 d}{\rho\kappa} \quad (12)$$

We assume free boundaries as appropriate for a convecting layer within an atmosphere. The velocity boundary conditions are¹⁰

$$\mathbf{u} \cdot \mathbf{k} = \frac{d^2}{dz^2}(\mathbf{u} \cdot \mathbf{k}) = 0 \quad \text{at } z = \pm \frac{1}{2} \quad (13)$$

Two special cases illustrate the different roles of radiative transfer within the fluid and at the boundaries. In the limit $\gamma \rightarrow \infty$, equation (11) becomes the isothermal condition $\theta = 0$ commonly used in convection theory. This case of perfectly conducting boundaries neglects the role of radiation in exchanging heat at the boundaries. The effect of the τ^{-1} term, which represents radiative transfer within the fluid, is then to increase the stability of the fluid and to decrease the aspect ratio of convection cells. This result was also obtained by Goody¹³, who incorporated the equations of radiative transfer into the formulation of the convection problem but used the boundary condition $\theta = 0$. The stabilizing effect occurs because radiative

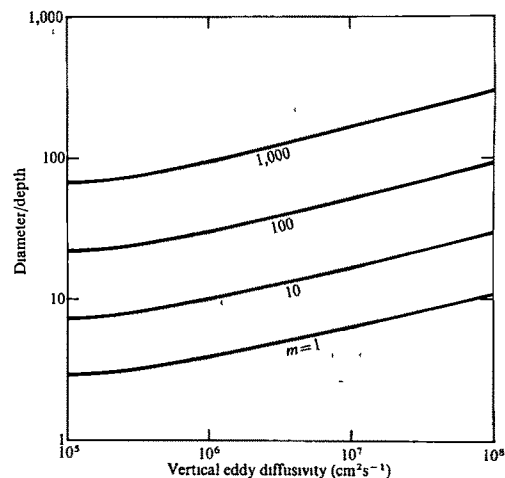


Fig. 3 Aspect ratio for the minimum Rayleigh number, as a function of κ , for various values of m .

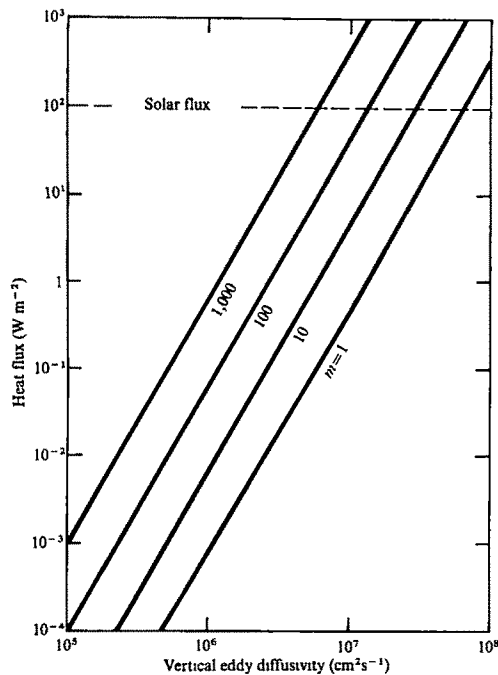


Fig. 4 Minimum convective heat flux required to sustain convection for the indicated values of κ and m

transfer adds to the effect of conduction in equalizing temperature differences between different parts of the fluid. The decrease in aspect ratio occurs because radiative transfer, unlike conduction, does not depend on the proximity of hot and cold fluid elements. In ordinary convection problems, cells with low aspect ratios are not seen because the proximity of hot upwelling regions and cold downwelling regions leads to large horizontal conductive heat transfer. When radiative transfer carries a significant amount of heat, this effect becomes less important and the cells tend to have lower aspect ratios.

The limit $\tau \rightarrow \infty$ accounts for radiative transfer at the boundaries but not within the bulk of the convecting fluid. As γ decreases the fluid becomes less stable and the aspect ratio of convection cells increases^{11,14,15}. The smaller γ is, the less efficient is radiation in exchanging heat at the boundaries. This explains the decreased stability of the radiative-convective static state. Fluid parcels must also spend more time at the boundaries exchanging heat, and this results in larger aspect ratios.

Thus, radiative transfer at the boundaries and within the fluid have opposite effects on the aspect ratio. In addition, anisotropic diffusivities with $m > 1$ tend to increase the aspect ratio^{16,17}. To see which effect 'wins' we solve the boundary value problem for realistic parameter values. Except for the eddy diffusivities, these are known at least in order of magnitude: $d \sim 10$ km, at 60 km altitude $c = 8.6 \times 10^6$ erg g⁻¹ K⁻¹, $\rho = 0.5 \times 10^{-3}$ g cm⁻³, and $T_0 = 270$ K from *in situ* probe measurements⁴, at cloud levels $\tau^* \sim 10^7$ s according to radiative transfer calculations⁷. Using these values, equations (10) and (12) give

$$\tau = \kappa / (10^5 \text{ cm}^2 \text{ s}^{-1}) \quad (14)$$

$$\gamma = (10^6 \text{ cm}^2 \text{ s}^{-1}) / \kappa \quad (15)$$

Thus, choosing the vertical thermal diffusivity κ determines τ and γ . The boundary value problem can then be solved by standard methods¹⁰. The onset of convection is steady for finite γ (ref. 11) and the additional effects of finite τ and $m \neq 1$ should not change this result. Thus we set $\partial/\partial t = 0$ in equations (5)–(7). We assume a horizontal wavenumber a and obtain an eigenvalue problem that determines R , the Rayleigh number for the onset of convection, as a function of a .

Figure 2 shows R plotted against the aspect ratio ($2\pi/q$) for different values of m , obtained by numerical solution of the eigenvalue problem with $\kappa = 10^6 \text{ cm}^2 \text{ s}^{-1}$. To match the cell dimensions observed in the Venus atmosphere, the minimum R must occur for aspect ratios of 10–100. Figure 2 shows that this occurs for $\kappa \sim 10^6 \text{ cm}^2 \text{ s}^{-1}$ and $m \sim 100$. These are magnitudes commonly used for the terrestrial atmosphere^{16–18}. Thus our theory can explain the high aspect ratios using reasonable parameter values.

There are, of course, many combinations of κ and m which give the desired aspect ratios. Figure 3 plots the aspect ratio at minimum R as a function of κ . There is a trade-off between m and κ : a lower eddy anisotropy is allowed if the magnitude of the eddy diffusivity is larger. For example, $m \sim 10$ and $\kappa \sim 10^7 \text{ cm}^2 \text{ s}^{-1}$ will also give aspect ratios in the range 10–100. Figure 3 also shows the influence of radiative transfer on the aspect ratio. For $\kappa \leq 10^5 \text{ cm}^2 \text{ s}^{-1}$ the aspect ratios approach the values one would obtain by neglecting radiative transfer ($\gamma = \tau = \infty$). For larger values of κ there is a noticeable effect: the boundary radiation effect (γ) dominates the internal radiation effect (τ), so that the aspect ratio is increased. Larger values of κ imply that eddies transfer heat very efficiently within the fluid. At the boundaries, heat is transferred by radiation because eddies are suppressed in the stable layers which bound the convecting region. This radiative transport is comparatively inefficient, causing larger aspect ratios as explained above.

The heat transport due to convection provides an important constraint on the size of κ . For instance, using $\kappa = 10^6 \text{ cm}^2 \text{ s}^{-1}$ and $m = 100$ we find a minimum Rayleigh number of 2.27×10^4 (Fig. 2). Using $\nu = \kappa$ we find

$$|dT_0/dz^*| - \Gamma_{ad} = \nu \kappa R / \alpha g d^4 = 7 \times 10^{-4} \text{ K km}^{-1} \quad (16)$$

This is the minimum superadiabaticity required to sustain convection in this case. It implies a minimum convective heat flux

$$\rho c \kappa (|dT_0/dz^*| - \Gamma_{ad}) = 6 \times 10^{-2} \text{ W m}^{-2} \quad (17)$$

Increasing κ will rapidly increase the heat flux to values greater than the downward solar radiative flux at cloud levels (estimated at 10^2 W m^{-2} from Pioneer Venus probe data¹⁹), a non-physical result. Figure 4, displaying the results of calculations analogous to equations (16) and (17) for different m and κ , shows that this limit is reached for $\kappa \geq 10^8$ when $m = 1$. Thus the combination $m = 1$, $\kappa \geq 10^8$ is ruled out even though it provides the correct aspect ratios (Fig. 3).

Conclusions

Our results show that the apparent flatness of large-scale convection cells in the Venus clouds can be explained by anisotropic eddy diffusion and radiative transfer effects. Horizontal eddy diffusivities must be at least 10-fold greater than vertical diffusivities. Anisotropy ratios $\geq 1,000$ are sufficient to account for the flattening in cases where the vertical eddy diffusivity is $\geq 10^5 \text{ cm}^2 \text{ s}^{-1}$ and the effects of radiative transfer are negligible. For vertical diffusivity $\geq 10^5 \text{ cm}^2 \text{ s}^{-1}$ radiative transfer contributes to the flattening. But radiative transfer alone cannot account for the apparent aspect ratios, this would require vertical heat diffusivities large enough to give a non-physically high value of convective heat transport.

The present theory is greatly simplified by the linearization of the equations about a static basic state. We have not only assumed small velocity perturbations but we have also made the Boussinesq approximation, which assumes small density variations. The latter is strictly valid only when the depth of the convecting layer is smaller than a scale height¹² (6 km in the region of interest). Finally, we have accounted for radiative heating within the convecting layer and have found that this effect tends to reduce the aspect ratio of the cells. Yet we cannot rule out the possibility that in fully developed convection the opposite effect occurs. In fact, Tritton has proposed on the basis of laboratory studies that internal heating contributes to the flattening of the convection cells on Venus²⁰. Considerations

such as this underscore the importance of studying nonlinear effects

We thank F H Busse for helpful discussions. This research was supported by NASA under NSG 7164 and NAS 2-9131

Received 23 September, accepted 22 December 1980

- 1 Murray, B C *et al Science* **183**, 1307-1315 (1974)
- 2 Rossow, W B, Del Genio, A D, Lumaye, S S, Travis, L D & Stone, P H *J geophys Res* **85**, 8107-8128 (1980)
- 3 Belton, M J S, Smith, G R, Schubert, G & Del Genio, A D *J Atmos Sci* **33**, 1394-1417 (1976)
- 4 Seiff, A *et al J geophys Res* **85**, 7903-7933 (1980)
- 5 Schubert, G *et al J geophys Res* **85**, 8007-8025 (1980)
- 6 Counsellman, C C, Gourevitch, S A, King, R W, Loriot, G B & Ginsberg, E S *J geophys Res* **85**, 8026-8030 (1980)
- 7 Pollack, J B & Young, R E *J Atmos Sci* **32**, 1025-1037 (1975)
- 8 Pollack, J B, Toon, O B & Boese, R J *J geophys Res* **85**, 8223-8231 (1980)
- 9 Agee, E M, Chen, T S & Dowell, K E *Bull Am met Soc* **54**, 1004-1012 (1973)
- 10 Chandrasekhar, S *Hydrodynamic and Hydromagnetic Stability* (Clarendon, Oxford, 1961)
- 11 Sasaki, Y *J met Soc Jap* **48**, 492-501 (1970)
- 12 Spiegel, E A & Veronis, G *Astrophys J* **131**, 442-447 (1960)
- 13 Goody, R M *J Fluid Mech* **1**, 424-435 (1956)
- 14 Sparrow, E M, Goldstein, R J & Jonsson, V K *J Fluid Mech* **18**, 513-528 (1964)
- 15 Jakeman, E *Phys Fluids* **11**, 10-14 (1968)
- 16 Priestly, C H B *Tellus* **14**, 123-124 (1962)
- 17 Ray, D *Tellus* **17**, 434-439 (1965)
- 18 Holton, J R *An Introduction to Dynamic Meteorology* 2nd edn (Academic, New York, 1979)
- 19 Tomasko, M G *et al J geophys Res* **85**, 8187-8199 (1980)
- 20 Tritton, D J *Nature* **257**, 110-112 (1975)

The structure of eight distinct cloned human leukocyte interferon cDNAs

David V. Goeddel, David W. Leung, Thomas J. Dull, Mitchell Gross, Richard M. Lawn, Russell McCandliss*, Peter H. Seeburg, Axel Ullrich, Elizabeth Yelverton & Patrick W. Gray

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

* Roche Institute of Molecular Biology, Nutley, New Jersey 07110, USA

Eight classes of human leukocyte interferon (LeIFN) cDNA clones have been identified in a cDNA library prepared from a myeloblastoid cell line. The nucleotide sequences demonstrate that the multiple human LeIFN genes code for a family of homologous, yet distinct proteins. One of the cDNA clones may have been derived from the transcription of a LeIFN pseudogene.

INTERFERONS are a family of proteins which confer viral resistance in target cells^{1,2} as well as inhibiting cell proliferation and modulating immune response (see refs 2, 3 for reviews). Human interferons were originally classified by their respective cells of origin. The major interferons produced in leukocytes (LeIFN or IFN- α) and in fibroblasts (FIFN or IFN- β) have been shown to be antigenically distinct, acid-stable molecules collectively designated 'classical' or type I interferons. Intense clinical interest in IFN coupled with the inability to obtain large amounts of the purified protein from human cells has spurred efforts to produce high yields of human interferon in the bacterium *Escherichia coli* by recombinant DNA procedures. Several groups have recently reported the bacterial synthesis of various interferons⁴⁻⁹, and the efficacy of such material in limited animal testing has been demonstrated⁵.

We recently described the construction of a collection of cDNA clones prepared using 12S RNA isolated from a human myeloblastoid cell line⁵. Several cDNA clones were identified by hybridization with synthetic DNA probes designed from partial amino acid sequence information of human leukocyte interferon. The cDNA insert of one of the hybrid plasmids (pL31) was sequenced and used to construct a plasmid which directed the synthesis of large amounts of a biologically active human LeIFN, termed LeIFN A (ref 5). Restriction endonuclease analysis showed extensive differences between the additional hybridizing clones, suggesting that several different human LeIFNs might be expressed by the myeloblastoid cell line KG-1 (ref 10). Furthermore, pL31 differs considerably from the human LeIFN cDNA clone described by Weissmann and colleagues⁴ in both nucleotide sequence and deduced amino acid sequence, yet both of the encoded LeIFNs show antiviral activity^{4,5}.

Additional evidence for the existence of a multigene family of human LeIFNs is found in the heterogeneity of the IFN proteins prepared from human leukocytes¹¹ and lymphoblasts^{12,13}. NH₂-terminal amino acid sequences¹²⁻¹⁴ and amino acid sequences of tryptic and chymotryptic peptides¹³ have revealed at least five

distinct, but homologous, LeIFNs. At least eight distinct chromosomal LeIFN genes have recently been isolated from a human gene library (ref 15, and R L and A U, unpublished results). In contrast, there is no evidence for multiple human FIFN genes. NH₂-terminal amino acid sequences of FIFN determined by two different groups^{16,17} agree with results obtained from nucleotide sequencing of cloned FIFN cDNAs^{7,18,19} and FIFN mRNA^{20,21} in suggesting that there may be only one gene for FIFN. We describe here genomic hybridization data consistent with the existence of about ten human LeIFN genes and a single FIFN gene, and present the DNA sequences of eight distinct LeIFN cDNA clones identified in a cDNA library prepared from the myeloblastoid cell line KG-1.

Evidence for multiple LeIFN genes

To estimate the number of LeIFN genes in the human genome, human spleen DNA was digested with various restriction endonucleases, electrophoresed through agarose gels and transferred to nitrocellulose paper²². Hybridizations were done using radiolabelled LeIFN and FIFN⁷ cloned cDNAs (Fig 1). The number of LeIFN genes can be estimated from the number of hybridizing bands in the lanes containing *Hind*III and *Taq*I digests. As none of the LeIFN cloned cDNAs contains either of these restriction sites (see below), and because the human LeIFN genes do not seem to have any intervening sequences (ref 15, and R L and A U, unpublished results), it is likely that each band corresponds to at least one gene. Therefore, the hybridization patterns suggest the presence of at least ten LeIFN genes and only one FIFN gene. Identical patterns were obtained for the LeIFN genes regardless of whether only LeIFN A cDNA was used as the radiolabelled probe or several different types of LeIFN cDNAs were used as a mixed probe. The fibroblast interferon gene contains an internal *Bgl*II site^{7,18,19}. Only one hybridizing band was obtained for FIFN in the *Bgl*II digest because the hybridization probe was an isolated *Bgl*II-*Pvu*II fragment from cloned FIFN cDNA (ref 7).

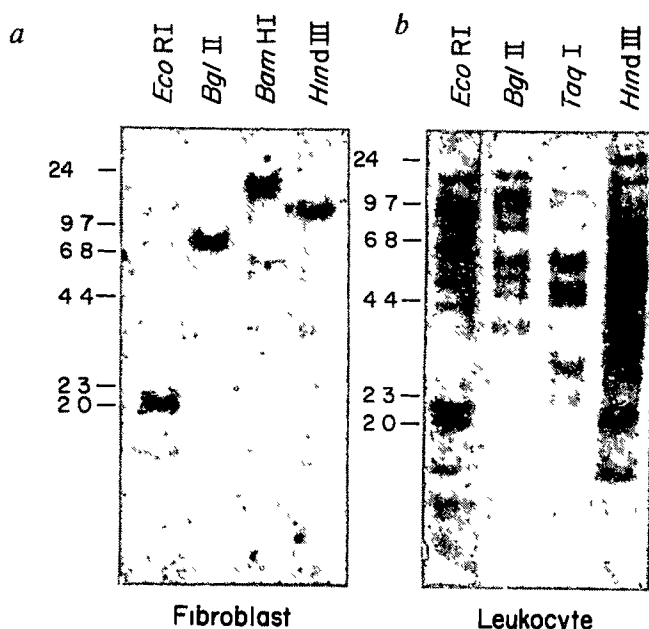


Fig. 1 Hybridization of human DNA with interferon cDNA probes. Human spleen DNA was isolated by the procedure of Blin and Stafford⁴³ and digested with the indicated restriction endonucleases. Samples (5 μ g) were electrophoresed in a horizontal 0.8% agarose slab gel and transferred to nitrocellulose filter paper²². Isolated DNA fragments consisting of *a*, a *Bgl*II-*Pvu*II fragment from cloned human IFN cDNA (ref. 7) or *b*, LeIFN A, B, C and E interferon cDNA sequences were ³²P-labelled to a specific activity of $\sim 10^6$ c.p.m. μ g⁻¹ by random calf thymus DNA fragment primed replication⁴⁴. Hybridizations were done using 5×10^6 c.p.m. of probe at 42 °C in 50% formamide, $5 \times$ SSC, $5 \times$ Denhardt's reagent⁴⁵, 50 mM sodium phosphate buffer, pH 6.5, 200 μ g ml⁻¹ denatured salmon sperm DNA and 10% dextran sulphate. After 24 h, the filters were washed at 60 °C in $0.1 \times$ SSC, 0.1% SDS as described elsewhere²⁷. Sizes in kilobase pairs (kb) of bacteriophage λ DNA digested with *Hind*III and included in the gels as size standards are indicated. All the LeIFN cDNAs described in this paper readily cross-hybridize so that hybridization patterns similar to those in *b* have been obtained with single LeIFN cDNA probes. Two additional *Bgl*II fragments of ~ 2.0 and 1.5 kilobase pairs hybridize to LeIFN cDNA probes but are not discernible in this figure.

Identification and classification of multiple LeIFN cDNA clones

To investigate the multiplicity of cloned human LeIFN cDNAs, an *in situ* colony screening procedure²³ was carried out on 4,000 individual colonies of our cDNA library. About 50 clones hybridized to a 260-base pair *Bgl*II fragment from pL31 encoding amino acids 61–150 of LeIFN A. Of these, 33/50 contained cDNA inserts of 700 base pairs or longer and were selected for further analysis. The cDNA inserts were mapped using the restriction enzymes *Pvu*II, *Bgl*II and *Eco*RI, because *Pvu*II and *Bgl*II had been found to cleave LeIFN A cDNA⁵ and

all three cut the LeIFN cDNA cloned by Nagata *et al.*⁴ These restriction endonuclease maps allowed the division of the LeIFN cDNAs into eight groups (A to H, see Fig. 2). The largest cDNA inserts in each group were sequenced by a combination of the Maxam–Gilbert chemical procedure²⁴ and the dideoxy-chain termination procedure using the bacteriophage M13 cloning vector mp7²⁵. It was possible that different LeIFN cDNAs would display the same restriction enzyme map. Because the 3'-noncoding regions of the eight sequenced cDNA clones showed much less homology than their respective coding portions, partial sequencing was done on the 3' regions of several additional cDNAs (five cDNAs each of LeIFN A and LeIFN D, and all the cDNAs in the remaining six groups). This analysis revealed that one of the cDNA clones originally classified as type C was actually a LeIFN H-type cDNA clone which did not extend to the single 5' *Pvu*II site, giving a total of two LeIFN H cDNA clones. However, as discussed in more detail below, the nucleotide sequences of these two LeIFN H cDNA clones differ at three positions.

The restriction enzyme analysis had shown that 14 of the 33 LeIFN cDNA clones were of type LeIFN A and 9 were LeIFN D. No variations in sequence were found among the five LeIFN A cDNAs or the five LeIFN D cDNAs which were partially sequenced—this suggests that the restriction maps are indicative of distinct classes of LeIFNs. LeIFNs A and D were the only groups represented more than twice in the collection of LeIFN cDNA clones. This result was not an artefact due to use of LeIFN A cDNA as the original hybridization probe; no additional LeIFN cDNA clones were found when the 4,000 colonies of the cDNA library were tested again using mixed LeIFN A, B, C, E and F cDNA probes.

The complete sequences of the longest cDNA inserts for each of the eight classes of LeIFN are shown in Fig. 3. Pairwise comparisons of these sequences show that homologies in the coding regions range from 84 (LeIFN B versus LeIFN E) to 94% (LeIFN C versus LeIFN F). The DNA sequences were used to predict the primary protein structure of the different human LeIFNs (Fig. 4). The cDNA clones for LeIFN types A, B, C, D, F and H contain complete coding sequences for the 23-amino acid precursor signal peptide^{4,5} and the 166 amino acids of the mature LeIFNs (165 amino acids for LeIFN A, which has a deletion at position 44). The 5' end of the LeIFN E cDNA begins in the signal peptide region, and also contains an extra base pair at position 187 (Fig. 3). LeIFN G is not a full-length cDNA copy of LeIFN mRNA and contains an incomplete coding sequence.

LeIFN D has the same restriction pattern as the LeIFN cDNA clone described by Nagata *et al.*⁴ The only difference between the two clones in predicted amino acid sequence is at position 114 where valine in LeIFN D replaces alanine²⁶. Weissmann and co-workers⁶ have also recently sequenced a partial-length LeIFN cDNA clone of a second type (which they named α_2) which differs from LeIFN A only at the codon for amino acid 23 (arginine in α_2 , lysine in LeIFN A).

Table 1 Pairwise comparisons of differences in coding sequences of human LeIFNs

	A	B	C	D	E	F	G	H	
A		31 (19)	31 (19)	29 (17)	41 (25)	30 (18)	18* (14)	28 (17)	Mature LeIFNs
B	7 (30)		32 (19)	38 (23)	45 (27)	31 (19)	23* (17)	29 (17)	
C	7 (30)	8 (35)		32 (19)	41 (25)	18 (11)	21* (16)	23 (14)	
D	6 (26)	9 (39)	8 (35)		44 (27)	28 (17)	17* (13)	32 (19)	
E	—	—	—	—		37 (22)	24* (18)	42 (25)	
F	7 (30)	8 (35)	0 (0)	8 (35)	—		15* (11)	28 (17)	
G	—	—	—	—	—	—		18* (14)	
H	6 (26)	6 (26)	6 (26)	4 (17)	—	6 (26)	—		
LeIFN signal peptides									

The number of amino acid replacements in each pair of coding sequences are shown. The 23 amino acid signal peptides are compared in the lower left part and the 166 amino acid mature LeIFNs are compared in the upper right part of the table. The total number of differences between pairs is listed first, followed by the percentage difference (numbers in parentheses). Signal sequences for LeIFNs E and G are not known, thus there are no values for comparisons involving either of these sequences.

* Only the 133 C-terminal amino acids of LeIFN G can be deduced from the cDNA sequence, thus in pairwise comparisons involving LeIFN G only differences in these 133 amino acids are considered.

2515 R 416

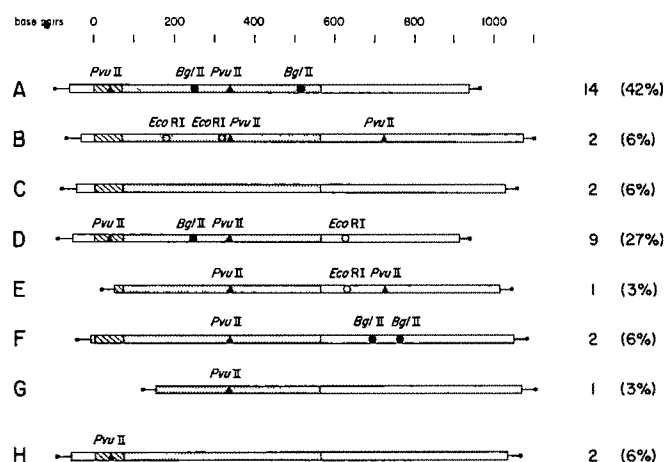


Fig. 2 Restriction endonuclease maps of the eight types of LeIFN cloned cDNAs (A-H). 33 LeIFN cDNA clones with *PstI* cDNA inserts >700 base pairs were grouped into these eight classes. The hybrid plasmids were constructed by the dC dG tailing method⁵. Therefore, the cDNA inserts could be excised using *PstI*. The lines at the end of each cDNA insert represent the flanking homopolymeric dC dG tails. The positions of *PvuII*, *EcoRI* and *BglII* restriction sites are indicated. Shaded regions indicate the coding sequences of mature LeIFNs, cross-hatched regions indicate signal peptide coding sequences, and the open regions show 3' and 5' noncoding sequences. The numbers at the right are the total number of cDNA clones in each group, with the per cent in each group in parentheses.

A LeIFN pseudogene is transcribed

The coding region of the LeIFN E cDNA clone contains one nucleotide more than the other LeIFN cDNAs (Fig. 3). The predicted amino acid sequence from this nucleotide insertion onwards (amino acid residue 40 in LeIFN E) bears no resemblance to the other LeIFN protein sequences and results in several in-phase termination codons, the first of which would appear at amino acid position 60. Therefore, the amino acid sequence is shown in Fig. 4 as if this insertion had not occurred. Even with this adjustment, termination codons are encountered at positions 102 and 155—thus they are found in all three reading frames. As a cDNA was cloned that contains an additional nucleotide in the coding sequence for a LeIFN, an artefact may have been created during the process of reverse transcription of mRNA and sequence amplification in *E. coli*. However, the presence of the two additional termination codons could imply that LeIFN E is a copy of a pseudogene analogous to the α - and β -globin pseudogenes²⁷⁻²⁹. If so, this is the first reported evidence of a pseudogene being transcribed into RNA. The biological significance of the transcription and possible translation of this pseudogene has not been determined. The fact that only one cDNA clone of a LeIFN pseudogene was obtained in our extensive search for LeIFN cDNA clones suggests that the transcription of such genes may be very rare. This observation is further supported by our recent isolation from a human chromosomal DNA library of an additional LeIFN pseudogene with a DNA sequence distinct from the eight LeIFN cDNA clones described here (unpublished results).

LeIFN signal peptides

The signal sequences of human LeIFNs A, B, C, D, F and H possess features common to all eukaryotic secretory protein signal sequences. Each contains a long stretch of hydrophobic amino acid residues, perhaps important for association with, and traversal of, the rough endoplasmic reticulum membrane. The hydrophobic regions are flanked by charged and more polar amino acid residues. The signal sequences are rich in serine residues proximal to the prepeptide cleavage site, analogous to the signal sequences of rat proalbumin³⁰ and certain mouse

immunoglobulin chains³¹. Within the cell, removal of the signal peptides is presumably achieved by peptide cleavage between glycine and cysteine residues in all of these LeIFNs.

The signal sequences of many different proteins have been determined and compared³². No major sequence homologies have been detected, the most general feature may be a common secondary or tertiary structure (see ref. 33 for review). It is possible that the junction of distinct structural domains within the signal peptide is recognized by a processing enzyme or signal peptidase³⁴, resulting in production of the mature secreted protein and complete translocation into the lumen of the rough endoplasmic reticulum.

The signal sequences of the LeIFNs reported here are only ~70% homologous when compared with each other, except for types C and F, which are identical (Table 1). Only 11 of the 23 amino acids (43%) occur in invariant positions in all six prepeptides. The sequence of the human FIFN signal peptide^{7,18-20} is completely different from those of the human LeIFNs.

Mature leukocyte interferons

The mature LeIFNs are more closely related in their amino acid sequences (~80% homology) than are their corresponding signal peptides (see Table 1). A total of 79 out of 166 amino acids (48%) of the mature polypeptide are identical in all sequences examined (Fig. 4). If the LeIFN E pseudogene amino acid sequence is ignored, then 99 amino acids (60%) occupy identical positions. When the human FIFN sequence^{7,18,19,21} is compared with the LeIFN sequences, only 39 amino acids out of 166 (23%) appear in identical positions (Fig. 4). The overall distribution of amino acid differences along the LeIFN molecules seems to be random, with the possible exception of the region between amino acids 115 and 151. All the LeIFN cDNA clones (again excluding LeIFN E) code for the same amino acid in 31 out of 37 (84%) positions in this region. Residues 116-150 also display the greatest similarity to FIFN (17 out of 35 positions identical), which suggests that this region is of functional importance.

The changes in amino acid sequences between the LeIFNs are largely due to single nucleotide changes, there are several sites (that is, positions 37, 107 and 160) where nonconservative amino acid changes occur. Note that the nucleotide sequence of LeIFN B contains an insertion of a thymidine residue after nucleotide 362 (codon 98), which changes the translational reading frame and results in four distinct codon changes. After position 373 (codon 101) an adenosine residue has been deleted which restores the original reading frame.

	97					102
	Glu	Ala	Cys	Val	Ile	Gln
LeIFN A	GAA	GCC	TGT	GTG	ATA	CAG
	Glu	Val	Leu	Cys	Asp	Gln
LeIFN B	GAA	GTC	CTG	TGT	GAT	CAG

The shorter (~800 base pairs) of the two LeIFN H-type cDNA clones, which we have designated LeIFN H₁, has a deletion (cytidine, nucleotide number 545) which is followed six nucleotides later by a correcting insertion (guanosine, after position 551).

	158				162
	Leu	Gln	Lys	Arg	Leu
LeIFN H	TTG	CAA	AAA	AGA	TTA
	Leu	Lys	Lys	Gly	Leu
LeIFN H ₁	TTG	AAA	AAA	GGA	TTA

Except for this difference and a single base change in the 3' noncoding regions, the nucleotide sequences of LeIFN H and LeIFN H₁ cDNAs are identical. It is possible that these two cDNA clones may represent allelic forms of a single LeIFN gene found in the KG-1 cell line.

The location of cysteine residues is often highly conserved in proteins. All the LeIFNs sequence code for cysteines at positions 1 and 139 of the mature protein. All predicted LeIFNs

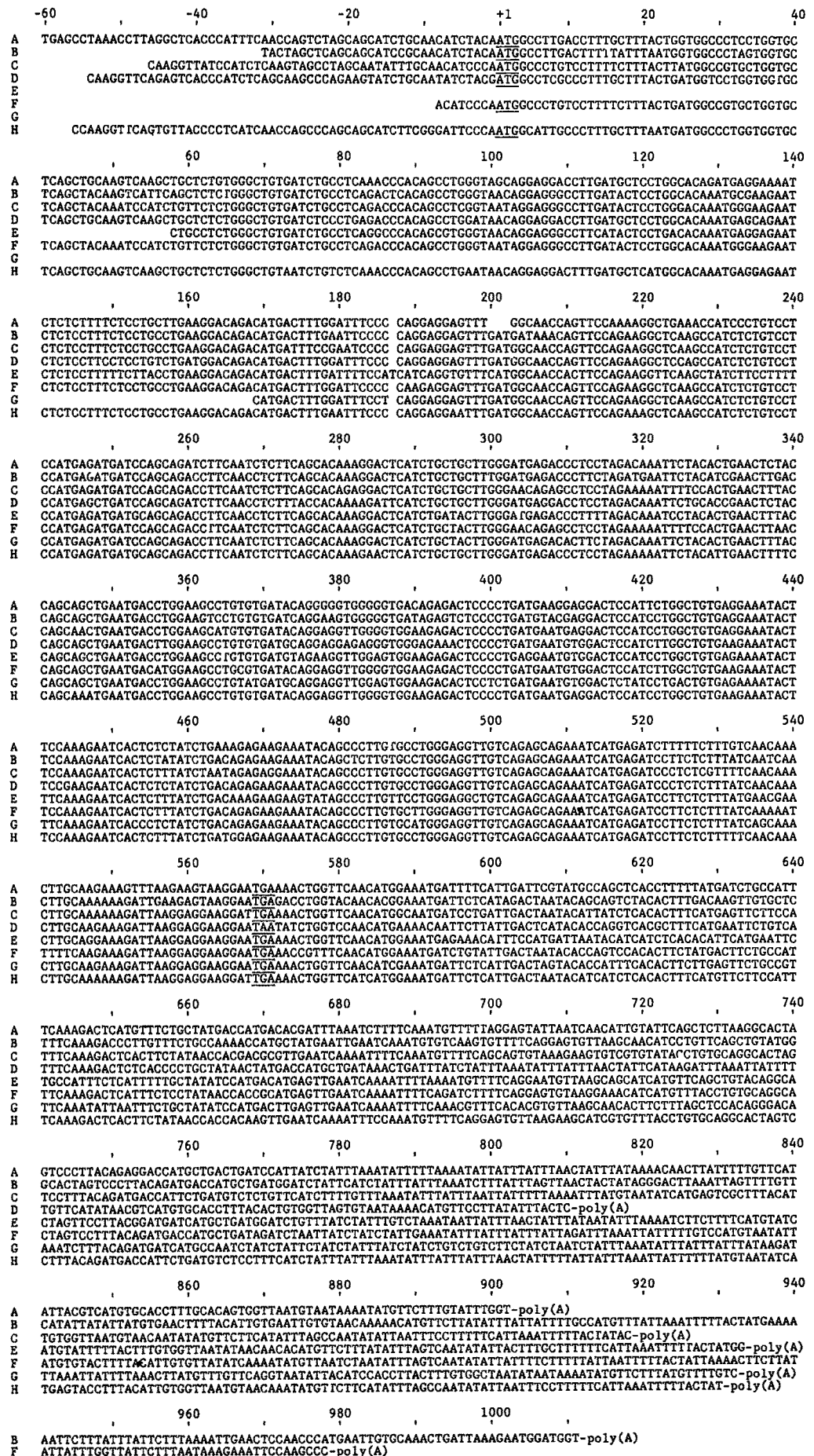


Fig. 3 Nucleotide sequences of the eight LeIFN cDNAs (A-H) DNA sequences of the *Pst*I inserts shown in Fig 2 were determined by the Maxam-Gilbert technique²⁴ and by the dideoxy chain termination procedure after subcloning into the M13 vector mp7 (ref 25). The ATG translational initiation codon and the termination triplet for each LeIFN are underlined. The gaps at position 187 were introduced to maximize homologies in the coding regions.

LeIFN ⁶¹	S10	S20	S23	10	20	30	40
A	MALTFALLVALLVLSCKSS	CSVGCGLPQTHSLGSRRTLMLLA	QMRKISLFSCLKDRHDFGFPQ				
B	MALTFYLLMAVLLVLSYKSS	CSLGCGLPQTHSLGNRRALILLGQ	MRISPFSSCLKDRHDFRIFPQ				
C	MALSFLLMAVLLVLSYKSS	CSLGCGLPQTHSLGNRRALILLGQ	MRISPFSSCLKDRHDFRIFPQ				
D	MALSFALLMAVLLVLSCKSS	CSLGCGLPQTHSLGNRRALILLGQ	MRISPFSSCLKDRHDFGFPQ				
E	MALSFLLMAVLLVLSYKSS	CSLGCGLPQTHSLGNRRALILLGQ	MRISPFSSCLKDRHDFGFPQ				
F	MALSFLLMAVLLVLSYKSS	CSLGCGLPQTHSLGNRRALILLGQ	MRISPFSSCLKDRHDFGFPQ				
G	MALSFLLMAVLLVLSYKSS	CSLGCGLPQTHSLGNRRALILLGQ	MRISPFSSCLKDRHDFGFPQ				
H	MALPFSLMHALVLLSCKSS	CSLGCGLPQTHSLGNRRALILLGQ	MRISPFSSCLKDRHDFGFPQ				
A11	MA F L	VLS KS	S GC L	THSL	RR L	QM	IS SCL DRHDE PQ
	50	60	70	80	90	100	
A	EEF-GNQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
B	EEFDDKQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
C	EEFDDKQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
D	EEFDDKQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
E	EEFDDKQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
F	EEFDDKQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
G	EEFDDKQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
H	EEFDDKQFQKAETIPVLHEMIQQIFNLFSTKDS	SAAWDETLLDKFYTELQQLNDLEACVIQG					
A11	EEFD QEQKA	I VLHE	QQ ENLF	T SSA	LL F	EL QQ	ND E Q
	110	120	130	140	150	160	166
A	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
B	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
C	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
D	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
E	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
F	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
G	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
H	VGVYTETPLMKEDSILAVRKYFQRITLYLKEKKYS	PCAWEVVRAEIMRSFSLSTNLQKRLRSKE					
A11	V PLM DSIL	V KYF	RITLYL	E KYS	CAWEVVR	AEIMRS	S S Q L K

Fig. 4 Comparison of the eight LeIFN protein sequences (A-H) predicted from nucleotide sequences. The one-letter abbreviations recommended by the IUPAC-IUB Commission on Biochemical Nomenclature⁴⁶ are used. A, alanine, C, cysteine, D, aspartic acid, E, glutamic acid, F, phenylalanine, G, glycine, H, histidine, I, isoleucine, K, lysine, L, leucine, M, methionine, N, asparagine, P, proline, Q, glutamine, R, arginine, S, serine, T, threonine, V, valine, W, tryptophan, and Y, tyrosine. The numbers refer to amino acid position (S refers to signal peptide). The dash in the 165 amino acid LeIFN A sequence at position 44 is introduced to align the LeIFN A sequence with the 166 amino acid sequences of the other LeIFNs. The LeIFN E sequence was determined by ignoring the extra nucleotide (position 187 of Fig. 3) in its coding region. Asterisks indicate in-phase termination codons. Amino acids common to all LeIFNs (excluding the pseudogene LeIFN E) are also shown. The underlined residues are amino acids which are also present in human IFN^{7,18,19,21}.

except LeIFN E contain a cysteine at position 29, and all except LeIFN B (Cys 100) contain a cysteine at position 99. At least one intramolecular disulphide bridge has been predicted on the basis of loss of interferon activity on incubation with reducing agents³⁵. Allen and Fantes¹³ have shown that the N-terminal cysteine is probably involved in a disulphide bond. Furthermore, the presence of two disulphide bridges in the LeIFN A produced in *E. coli*⁵ has been established by analysis of tryptic digests of the purified protein. Cys 1 is bonded to Cys 99 and Cys 29 is bonded to Cys 139 (ref. 47). The fact that cysteine residues are also predicted in human IFN^{7,18,19,21} in the positions corresponding to amino acids 29 and 139 of LeIFN suggests that this disulphide bond also occurs in IFN.

NH₂-terminal sequences of human LeIFN species isolated from lymphoblasts^{12,13} and leukocytes¹⁴ have been determined by protein microsequence analysis. The sequences determined by Zoon *et al.*¹² and Levy *et al.*¹⁴ both have NH₂-terminal serine residues. The sequences published by Allen and Fantes¹³ begin with cysteine residues. The seven full-length LeIFN cDNAs (A-F, H) all code for mature LeIFNs beginning with cysteine. Although it is possible that a LeIFN cDNA clone will be identified which encodes an N-terminal serine, it is more likely that the determination of serine arose artefactually through oxidation of cysteine during the initial protein microsequencing procedure.

Disregarding the discrepancies at amino acid position 1, the following comparisons can be made. The 32 NH₂-terminal amino acids of the lymphoblastoid IFN sequence of Zoon *et al.*¹² have the greatest homology with LeIFN F, with only position 26 (Pro in LeIFN F, Leu in lymphoblastoid IFN) differing between the two species. The 22 NH₂-terminal amino acids of LeIFN determined by Levy *et al.*¹⁴ show the most homology with LeIFN A. These differ only at position 11 (Asn in the Levy sequence compared with Ser for LeIFN A). Allen and Fantes¹³ have determined the sequences of tryptic and chymotryptic peptides of two forms of LeIFN: IFN- α A and IFN- α B¹³. IFN- α A comprises at least two different polypeptides of 165 amino acids with the same amino acid deletion at position 44 that is predicted by the LeIFN A cDNA sequence. 140 of the 142 amino acids which have been determined from the IFN- α A peptides correspond to the LeIFN A cDNA sequence presented

here, only Asp 84 and Pro 88 of IFN- α A are not predicted by the LeIFN A cDNA sequence. IFN- α B consists of at least three polypeptides of 166 amino acids each and the residues present at 161 positions have been determined¹³. Peptides of IFN- α B are most closely related to LeIFN D, with only Met 60, Thr 64 and Ala 114 not being predicted by the LeIFN D cDNA sequence. The 77 N-terminal amino acids of LeIFN F are found in the IFN- α B peptide sequences. However, there are 11 differences between these two interferon sequences in the C-terminal portion (residues 78-166) of the polypeptides. Amino acid sequences deduced from LeIFN B, C, G and H cDNAs do not seem to be represented in either IFN- α A or IFN- α B peptide sequences. These comparisons suggest the presence of additional LeIFN genes, cDNA copies of which are not present, or have not been identified, in our cDNA library. Whether some of the slight differences between LeIFN sequences described above could be due to allelic variation in the human population awaits a detailed genomic characterization of the LeIFN gene family from an individual.

Noncoding sequences of LeIFN

A limited amount of sequence information is available for the 5' noncoding regions of the LeIFN cDNAs (Fig. 3). Obvious homologies exist in the regions immediately preceding the translational initiation codon, LeIFNs A, B and D are similar in this region, as are LeIFNs C and F.

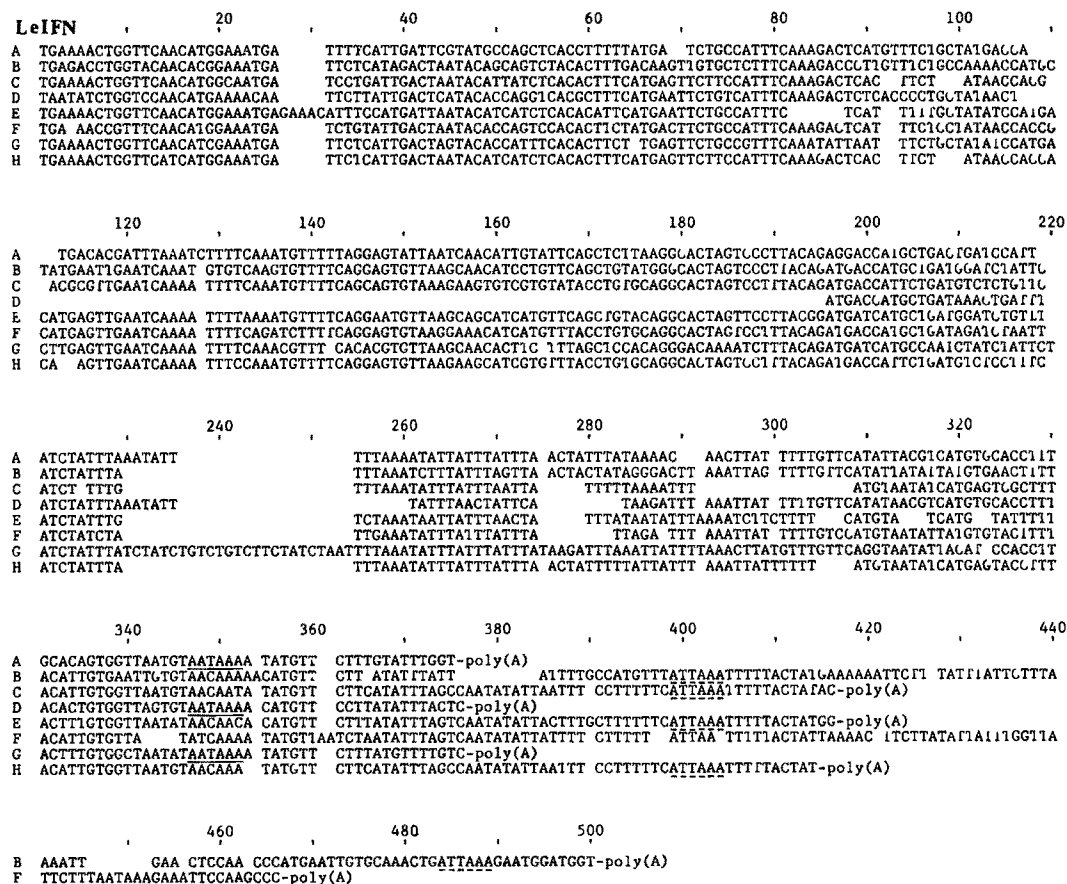
The 3' noncoding regions of the LeIFN cDNA clones vary in size from 242 (LeIFN D) to 441 (LeIFN B) base pairs, counting from the first nucleotide of the stop codon to the site of polyadenylation. Figure 5 shows an alignment of these sequences with gaps introduced to achieve maximum homology. These gaps may be sites of deletions or insertions occurring during the evolution of this multigene family³⁶. The overall homology between 3' noncoding regions is lower than that within the protein coding regions, although certain portions of the 3' noncoding region are relatively conserved. Both the reduced level of 3' noncoding homology and the substantial deletions in the sequence alignments resemble the situation found when comparing the analogous DNA sequences of members of the human β -like globin gene family³⁶. Such results suggest that the precise sequence of much of the 3' noncoding regions of

eukaryotic mRNAs is not essential for function. Experiments involving *in vitro* translation of mRNAs with large 3' deletions support this suggestion^{37,38}.

Efstratiadis *et al.*³⁶ have noted that short (2–8 base pairs) tandem repeats often surround the deletion sites in the β -like globin gene family 3' noncoding sequences. They suggested a mechanism of 'slippage' during DNA replication followed by scission of resulting single-stranded DNA loops as a possible cause of these deletion events during evolution. Such short tandem repeats also occur at many of the deletion points in the LeIFN cDNA sequences.

The hexanucleotide AATAAA precedes the site of polyadenylation in many eukaryotic cellular mRNAs by 15–25 bases³⁹. Although exceptions to the rule have been found in several RNA viruses⁴⁰, this sequence is thought to be involved in mRNA processing and/or polyadenylation. A variant of this sequence, AATTAAA, has recently been reported which precedes the poly(A) site in rat amylase⁴¹ and anglerfish somatostatin⁴² mRNAs. LeIFN A, D, F and G cDNAs contain the hexanucleotide AATAAA ending 20, 21, 14 and 21 nucleotides, respectively, from their polyadenylation sites. LeIFN B, C, E and H cDNAs do not contain this hexanucleotide. At the nucleotide position (nucleotides 347–352, Fig. 5) corresponding to AATAAA in the LeIFN A, D and G cDNAs, the sequence of the other cDNAs differ by one or two nucleotides. This sequence divergence may have led to the extension of the 3' noncoding region in those genes. While LeIFN F mRNA continues until an AATAAA is encountered 14 nucleotides before its poly(A) site, LeIFNs C, E and H possess the related hexanucleotide ATTAATA near their 3' termini (positions 399–404). LeIFN B also has the ATTAATA sequence at this position, but is not polyadenylated until another ATTAATA sequence is reached at positions 484–489. Note that the two cDNA clones most often isolated (LeIFN A and D) have the shortest 3' noncoding sequences and possess the canonical hexanucleotide at the same position.

Fig. 5 Alignment of the 3' noncoding sequences of eight human LeIFN cDNAs. The alignment of sequences between the termination codon and the poly(A) addition site was made using considerations described elsewhere³⁶. Blank areas indicate gaps introduced to achieve consistent alignments. The sequence AATAAA common to most eukaryotic polyadenylated mRNAs³⁹ is underlined. The related sequence ATTAATA, which may also serve as a polyadenylation signal, is indicated by dashed lines.



Significance of the LeIFN multigene family

We have presented here the sequences of eight distinct human LeIFN cDNAs, seven of which probably code for functional interferon polypeptides. LeIFN amino acid sequence information^{12–14} suggests the presence of additional LeIFN structural genes. The genomic blots shown in Fig. 1 and the isolation of chromosomal LeIFN sequences¹⁵ also support the concept of a LeIFN multigene family comprising at least 8–10 distinct human LeIFN-related genes, although some of these could be pseudogenes similar to LeIFN E.

The cDNA clones LeIFN A and D occur relatively frequently in our cDNA library⁵, which was prepared using RNA from the virus-induced myeloblastoid cell line KG-1¹⁰. In contrast, a major LeIFN species found in the lymphoblastoid Namalwa cell line¹² most closely resembles the polypeptide encoded by LeIFN F cDNA—this supports the suggestion of Nagata *et al.*¹⁵ that different cell types might preferentially express different LeIFN genes. Preferential expression of different IFNs might also occur in the same cell line in response to different viruses.

These multiple LeIFNs probably show distinct functional properties. Streuli *et al.*⁶ have shown that two different LeIFNs (IFN- α 1 and IFN- α 2) have striking differences in their antiviral activities on cells of different species. We have recently obtained expression in *E. coli* of the mature LeIFN polypeptides encoded by LeIFN A, B, C, D and F cDNAs (unpublished results) and all five show different cross-species specificities. The availability of these cloned LeIFNs should allow the assignment of the various activities characteristic of mixed leukocyte interferon preparations² to individual LeIFN species. A detailed understanding of the correlation between LeIFN structure and activity may lead to the design of new interferons highly active against particular diseases.

We thank Robert Swanson and Giuseppe Miozzari for their encouragement, and Jeanne Arch, Alane Gray and Suzanne Pfeffer for help in preparing the manuscript. This work was supported by Hoffmann-La Roche Inc. and Genentech Inc.

Received 27 October 1980, accepted 15 January 1981

- 1 Isaacs, A & Lindenmann, J *Proc R Soc B* **147**, 258-267 (1957)
- 2 Stewart, W E II *The Interferon System* (Springer, New York, 1979)
- 3 Baron, S & Dianzani, F (eds) *Tex Rep Biol Med* **35**, 1-573 (1977)
- 4 Nagata, S *et al Nature* **284**, 316-320 (1980)
- 5 Goeddel, D V *et al Nature* **287**, 411-416 (1980)
- 6 Streuli, M, Nagata, S & Weissmann, C *Science* **209**, 1343-1347 (1980)
- 7 Goeddel, D V *et al Nucleic Acids Res* **8**, 4057-4074 (1980)
- 8 Derynck, R *et al Nature* **287**, 193-197 (1980)
- 9 Taniguchi, T *et al Proc natn Acad Sci U S A* **77**, 5230-5233 (1980)
- 10 Koeffler, H P & Golde, D W *Science* **200**, 1153-1154 (1978)
- 11 Rubinstein, M *et al Proc natn Acad Sci U S A* **76**, 640-644 (1979)
- 12 Zoon, K C *et al Science* **207**, 527-528 (1980)
- 13 Allen, G & Fantes, K H *Nature* **287**, 408-411 (1980)
- 14 Levy, W P, Shively, J, Rubinstein, M, Del Valle, U & Pestka, S *Proc natn Acad Sci U S A* **77**, 5102-5104 (1980)
- 15 Nagata, S, Mantei, N & Weissmann, C *Nature* **287**, 401-408 (1980)
- 16 Knight, E Jr, Hunkapiller, M W, Korant, B D, Hardy, R W F & Hood, L E *Science* **207**, 525-526 (1980)
- 17 Stein, S *et al Proc natn Acad Sci U S A* (in the press)
- 18 Taniguchi, T, Ohno, S, Fujii-Kuriyama & Muramatsu, M *Gene* **10**, 11-15 (1980)
- 19 Derynck, R *et al Nature* **285**, 542-547 (1980)
- 20 Houghton, M *et al Nucleic Acids Res* **8**, 1913-1931 (1980)
- 21 Houghton, M *et al Nucleic Acids Res* **8**, 2885-2894
- 22 Southern, E M *J molec Biol* **98**, 503-517 (1975)
- 23 Grunstein, M & Hogness, D S *Proc natn Acad Sci U S A* **72**, 3961-3965 (1975)
- 24 Maxam, A M & Gilbert, W *Meth Enzym* **65**, 499-560 (1980)
- 25 Messing, J, Crea, R & Seeburg, P H *Nucleic Acids Res* (in the press)
- 26 Mantei, N *et al Gene* **10**, 1-10 (1980)
- 27 Fritsch, E F, Lawn, R M & Maniatis, T *Cell* **19**, 959-972 (1980)
- 28 Proudfoot, N J & Maniatis, T *Cell* **21**, 537-544 (1980)
- 29 Proudfoot, N J, Shander, M H M, Manley, J L, Gelfer, M L & Maniatis, T *Science* **209**, 1329-1336 (1980)
- 30 Strauss, A W, Bennett, C D, Donohue, A M, Rodkey, J A & Alberts, A W *J biol Chem* **252**, 6846-6855 (1977)
- 31 Burnstein, Y, Zemell, R, Cantor, F & Schechter, I *Proc natn Acad Sci U S A* **74**, 3157-3161 (1977)
- 32 Habener, J F *et al Proc natn Acad Sci U S A* **75**, 2616-2620 (1978)
- 33 Blobel, G *et al Symp Soc exp Med* **33**, 9-36 (1979)
- 34 Jackson, R C & Blobel, G *Proc natn Acad Sci U S A* **74**, 5598-5602 (1977)
- 35 Mogensen, K E & Cantell, K *J gen Virol* **22**, 95-103 (1974)
- 36 Efstratiadis, A *et al Cell* **21**, 653-668 (1980)
- 37 Kronenberg, H M, Roberts, B E & Efstratiadis, A *Nucleic Acids Res* **6**, 153-166 (1979)
- 38 Nunberg, J H, Kaufman, R J, Chang, A C Y, Cohen, S N & Schumke, R T *Cell* **19**, 355-364 (1980)
- 39 Proudfoot, N J & Brownlee, G G *Nature* **263**, 211-214 (1976)
- 40 Porter, A G *et al Nature* **276**, 298-301 (1978)
- 41 MacDonald, R J *et al Nature* **287**, 117-122 (1980)
- 42 Hobart, P, Crawford, R, Shen, L P, Pictet, R & Rutter, W J *Nature* **288**, 137-141 (1980)
- 43 Blin, N & Stafford, D W *Nucleic Acids Res* **3**, 2303-2308 (1976)
- 44 Taylor, J M, Illemensee, R & Summer, S *Biochim biophys Acta* **442**, 324-330 (1976)
- 45 Denhardt, D T *Biochem biophys Res Commun* **23**, 461-467 (1966)
- 46 Dayoff, M O *Atlas of Protein Sequence and Structure* 5 Suppl 3 (National Biomedical Research Foundation, Washington DC, 1978)
- 47 Wetzel, R *Nature* **289**, 606-607 (1981)

Homology and concerted evolution at the $\alpha 1$ and $\alpha 2$ loci of human α -globin

Stephen A. Liebhaber, Michel Goossens & Yuet Wai Kan

Howard Hughes Medical Institute Laboratory and Department of Medicine, University of California, San Francisco, California 94143, USA

The identical structure of two racially distinct $\alpha 1$ -globin alleles and the high degree of homology between the $\alpha 1$ - and $\alpha 2$ -globin loci indicate that mechanisms exist for suppression of allelic polymorphisms and for exchange of genetic information within the α -globin gene complex.

THE human α -globin polypeptide chains are encoded at two closely linked loci^{1,2} on chromosome 16 (ref 3). Genomic DNA containing both of these genes has been cloned and mapped⁴. The loci are 3.2 kilobases apart and are oriented in the same transcriptional direction. The more 5' or leftward α -globin gene has been designated as $\alpha 2$ and the more 3' or rightward α -globin gene as $\alpha 1$. The α -globin proteins encoded at these loci have remained identical⁵ even though the duplication of the α -globin gene probably occurred before the mammalian divergence 85 Myr ago and possibly before the bird-mammal divergence a minimum of 300 Myr ago⁶. As gradual accumulation of mutations within these two genes would have caused divergence in the primary structure of the coded proteins over this time period, a mechanism for correcting one gene against another has been postulated⁷. The parallel or coincidental evolution of two genes (concerted evolution⁸) may occur by homologous but unequal cross-over between the two mutually correcting genes. Comparison of α -globin protein structure in primates indicates that this interaction between the two α -globin loci occurs at a high frequency⁶. Comparisons of the structure of the two γ -globin loci (γ^A and γ^G) support the concept of concerted evolution and suggest mechanisms for its implementation^{9,9}. Heteroduplex analysis⁴ and partial sequencing¹⁰ of the DNA within the human α -globin cluster have identified three homologous units that may align the $\alpha 1$ - and $\alpha 2$ -globin genes on sister chromatids, facilitating unequal cross-over between these two loci. To expand on these findings, we have sequenced the $\alpha 1$ -globin genes from two individuals of different races and compared these alleles with each other and with the structure of the $\alpha 2$ -globin gene¹¹. The two $\alpha 1$ -globin alleles are identical, whereas the $\alpha 1$ and $\alpha 2$ loci have diverged in a defined manner. These observations are consistent with frequent exchanges of DNA within the α -globin gene cluster.

Primary structure of two $\alpha 1$ -globin alleles is identical

The $\alpha 1$ -globin genes cloned from the DNA of an Asian (pGL $\alpha 1$) and a Caucasian (pRB $\alpha 1$) were sequenced in parallel. The sequencing scheme used is shown in Fig 1 and the regions

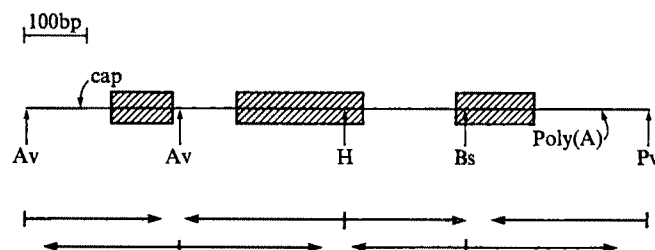


Fig. 1 Strategy for sequencing the $\alpha 1$ -globin genes. Vertical arrows denote restriction enzyme cleavage sites, *AvaI* (Av), *HindIII* (H), *BstE2* (Bs), *PvuII* (Pv). The positions of the capping and poly(A) addition sites as determined from the mRNA^{28,29} are shown. DNA was cut at each of the indicated restriction sites and either 5'-labelled with polynucleotide kinase (BRL) and [γ -³²P]ATP (Amersham, 3,000 Ci mmol⁻¹) after dephosphorylation of the DNA with bacterial alkaline phosphatase (BRL), or 3'-labelled with *Escherichia coli* DNA polymerase I, Klenow fragment (BRL) and the appropriate [α -³²P]dNTP (3,000 Ci mmol⁻¹). These labelled fragments were purified on 5% acrylamide gels and either strand separated³⁰ or secondarily cut by another restriction endonuclease. Gel-purified end-labelled pieces were sequenced by the method of Maxam and Gilbert³⁰. The horizontal arrows are proportional to the distance each fragment was read. All fragments were sequenced twice and in most cases on both strands. bp, Base pairs.

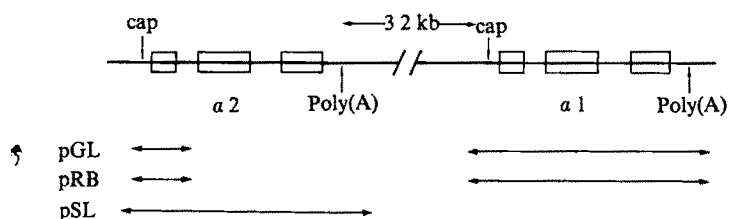


Fig. 2 Sequencing summary. The complete sequence of the $\alpha 2$ -globin allele pSL $\alpha 2$ has been reported¹¹. pGL $\alpha 1$ and $\alpha 2$ are pBR 322 subclones containing respectively the $\alpha 1$ - and $\alpha 2$ -globin genes from a single chromosome (M. Goossens, unpublished data). These were subcloned from a Charon 4A recombinant containing a 14.5-kilobase (kb) *Bam*HI fragment that includes both α -globin genes¹. The DNA originated from the leukocytes of a Chinese subject. pRB $\alpha 1$ and $\alpha 2$ are pBR 322 subclones containing the $\alpha 1$ - and $\alpha 2$ -globin genes from a single chromosome (γ H α G2)⁴ of a Caucasian patient and were provided by T. Maniatis. The plasmid subclones used for sequencing were grown in 15-l batches, amplified at A_{640} 0.7 with chloramphenicol ($150 \mu\text{g ml}^{-1}$), collected 18 h later and purified from a clarified lysate by CsCl isopycnic centrifugation. The extent to which each of these genes was sequenced is indicated by the horizontal arrows (see text for details). The $\alpha 1$ - and $\alpha 2$ alleles were sequenced in parallel to enhance their comparison and to minimize misreadings. Each gene was labelled at the same restriction site, chemical reactions were done simultaneously and, in most instances, the modified and cleaved fragments were electrophoresed side by side on the same gels.

compared in each clone are summarized in Fig. 2. No differences are present in the primary sequence as determined by a direct visual comparison of the adjacent sets of sequencing lanes. Likewise, the sequences of the 230-base pair *Ava*I fragments (Figs 1, 2) encompassing the 5' end of three $\alpha 2$ -globin genes (pSL $\alpha 2$, pGL $\alpha 2$ and pRB $\alpha 2$) are identical and limited sequencing of the 3' end of a fourth $\alpha 2$ -globin allele (Chinese) from the *Bst*E-2 site at AA104 (Fig. 1) to the polyadenylation [poly(A)] site (G. Temple, personal communication) is identical with our reported $\alpha 2$ sequence for a Caucasian (pSL $\alpha 2$)¹¹ (see Fig. 2). This degree of homology was unexpected as Jeffreys has estimated that 1% of the bases in the human β -globin cluster are sites of polymorphism¹². However, this 1% estimate is based on the identification of three restriction site polymorphisms located in the large intervening sequence of the γ and δ genes. Subsequent sequencing of these genes has revealed that an unusually high rate of sequence divergence exists in this region¹³ and that the two polymorphisms noted in the γ -globin genes are probably due to a single initial mutation in one of the γ -globin genes with subsequent transfer of this mutation to the other γ -globin gene⁹. Therefore a reliable overall rate of sequence divergence in the β -globin complex must await comparative sequencing along the entire length of the genes in question.

3' border of homology between the α -globin loci is near the UAA termination codon

Comparison of the α -globin gene sequence at the $\alpha 1$ and $\alpha 2$ loci¹¹ reveals no differences in the 5'-flanking region, the 5'-nontranslated segment, the first exon, or the first intervening sequence (IVS). Exons 2 and 3 each contain one silent third-base substitution. Codon 54 (Gln) is encoded by CAG in $\alpha 1$ and by CAA in $\alpha 2$, and codon 123 (Ala) by GCC in $\alpha 1$ and GCT in $\alpha 2$.

The IVS2 of the $\alpha 1$ -globin gene is nine bases longer and differs by three bases from the same region of the $\alpha 2$ -globin gene (pSL $\alpha 2$). Position 55 of IVS2 is T in $\alpha 2$ -globin and G in $\alpha 1$ -globin. The remaining two base changes and all nine inserted/deleted bases are clustered toward the 3' end of this intron. This finding is consistent with the previous comparison of the pRB $\alpha 1$ and pRB $\alpha 2$ restriction maps which identified a

seven-base insertion/deletion in the same region of IVS2 (ref. 4). Figure 3 shows an alignment of this region that maximizes homology and assumes that the nine bases were deleted as three separate events. The seven-base deletion is flanked by a four-base direct repeat. Deletions in prokaryotic models¹⁴ as well as a number of sequence deletion within the β -globin gene cluster¹³ are similarly flanked by direct repeats in the precursor molecule. Whether the underlined four-base repeat (GGCC, Fig. 3) results in deletion by strand mismatch during DNA replication as suggested by these authors remains speculative, as the current data do not eliminate the possibility that the differences were created by insertions or insertion/deletion combinations.

Heteroduplex mapping studies define several large repeated homologous regions encompassing the $\alpha 1$ - and $\alpha 2$ -globin genes⁴ (Fig. 4a). Proudfoot and Maniatis¹⁰ have sequenced the actual borders of some of these segments. The most leftward homology box begins just 3' to the poly(A) addition site of the pseudo- $\alpha 1$ - and $\alpha 2$ -globin genes. This area includes a large intergenic region of homology between pseudo- α and $\alpha 2$ and $\alpha 2$ and $\alpha 1$ (the X box in ref. 11) (Fig. 4a). Based on the assumption that the 3'-noncoding region of $\alpha 1$ - and $\alpha 2$ -globin mRNA was identical (see below), the poly(A) addition site of these two genes was thought to define the 3' border of a second set of homology regions (the Z box) which encompasses the two α -globin genes and extends 1.5 kilobases in the 5' direction. Our sequence comparison of the 3'-noncoding region of the $\alpha 1$ and $\alpha 2$ genes (Fig. 4b) defines 18 base differences plus a single-base insertion or deletion, making this region only 84% homologous. Thus we would reassign the 3' border of the Z box (the border of α -globin gene recombination—see below) to a position 14 bases 3' to the UAA termination codon. The assumption that the 3'-noncoding regions of both α -globin genes were identical was based on the determination of a unique sequence for this region from reticulocyte mRNA¹⁵. That sequence is almost identical to the gene sequence of the $\alpha 2$ -globin locus, whereas it differs from the $\alpha 1$ -globin gene at 18 bases and one insertion or deletion. Although a definite sequence assignment could not be made at seven positions, only three of these correlate with the base differences between the $\alpha 1$ - and $\alpha 2$ -globin genes. This disparity between the unique sequence of reticulocyte α -globin mRNA and the above genomic sequence data may be reconciled by suggesting that the mRNA transcribed from the $\alpha 2$ -globin locus predominates in circulating reticulocytes due to unequal production or degradation¹⁶.

The sequence divergence between $\alpha 1$ - and $\alpha 2$ -globin genes which begins 3' to the UAA terminator codon and extends into the 3'-flanking region is interrupted by a 50-base pair homologous stretch bridging the poly(A) site and the AAUAAA hexanucleotide (poly(A) signal)¹⁷ (Figs 4b, 5). This region (with the exception of the conserved hexanucleotide AAUAAA) is highly divergent in eukaryotic mRNA in general, and in β -globin can be deleted or inactivated without adversely affecting *in vitro* translational activity¹⁸. Among β -globin and β -like globin genes homology in this region is significant only for the hexanucleotide AAUAAA, its immediately surrounding bases and the sequences adjacent to the poly(A) addition site¹³. Thus the observed maintenance of perfect homology in a segment of

GGC_GCGG_CT_TGGGCC_____GCACT	$\alpha 2$
GGCTGCGGGCCTGGGCC_____CCACT	Possible transitional sequence
GGCTGCGGGCCTGGGCCCTCGGCCCACT	$\alpha 1$

Fig. 3 Deletion sites in the 3' end of IVS2. This sequence begins at base 101 of IVS2 and extends through the divergent region to a position 19 base pairs from the junction of IVS2 with exon 3. The sequence of the $\alpha 1$ -globin gene, the $\alpha 2$ -globin gene¹¹ and a possible intermediate sequence are aligned for maximal homology. The open spaces in the $\alpha 2$ -globin gene and the intermediate sequence represent deletions in this region relative to the $\alpha 1$ -globin gene. The dot indicates the position of a base change. The underlined tetranucleotide GGCC is discussed in the text.

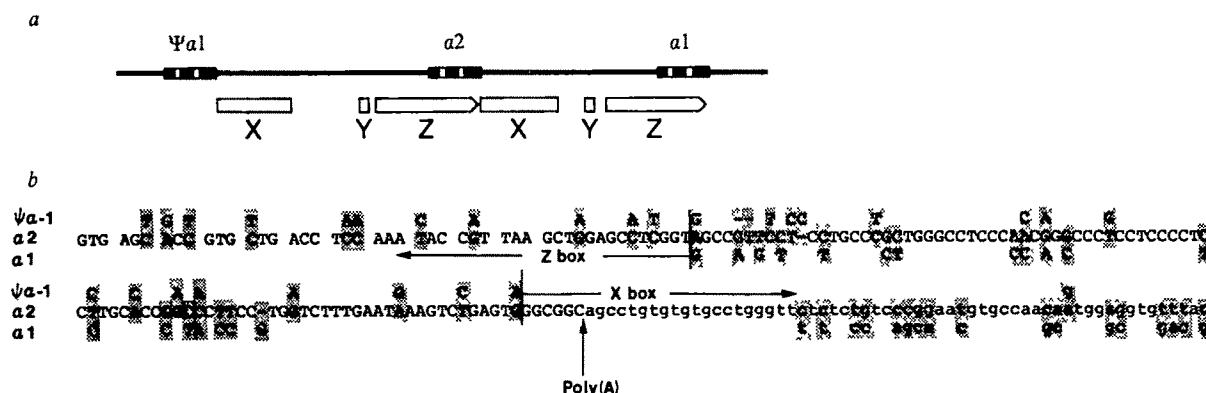


Fig. 4 Homologous regions within the segment of α -globin gene cluster containing the $\psi\alpha 1$ -, $\alpha 2$ - and $\alpha 1$ -globin genes. The rectangular blocks represent the respective genes and the open spaces within these blocks represent the introns. Regions that show identity by heteroduplex analysis⁴ or partial sequence determination¹⁰ are given the same letters (X, Y or Z). This figure is adapted from ref. 10 by permission of the author (N. J. P.). **b**, Comparison of the 3' ends of the $\alpha 1$ -globin, $\alpha 2$ -globin and $\psi\alpha 1$ -globin genes¹⁰. The sequence begins at the codon for AA132 of the functional α -globin genes and extends through the TAA ochre codon, the 3'-nontranslated region, and into the 3'-flanking region (lower-case letters). Base differences between the $\alpha 1$ -globin gene and the $\alpha 2$ -globin gene or $\psi\alpha 1$ -globin gene¹⁰ are indicated by the shaded boxes. Deletions are indicated by a hyphen. The region of homology encompassing the $\alpha 1$ - and $\alpha 2$ -globin genes (Z box) has its 3' border approximately 14 bases 3' to the TAA termination codon and the region of homology 3' to the $\psi\alpha 1$ - and $\alpha 2$ -globin genes (X box) has its 5' border six bases 5' to the poly (A) addition site.

this region is probably not entirely explained by functional constraints and could instead be due to a localized transfer of DNA between the two α -globin loci (see below).

Evolution of the α -globin genes

The $\alpha 1$ - and $\alpha 2$ -globin genes may exchange DNA by interaction between malaligned chromosomes. Such interactions may be facilitated by and result in large homologous regions within the α -globin complex (refs 4, 10, see also Fig. 4a). The degree of divergence within these regions would be proportional to the length of time since this exchange occurred. Divergence outside this unit (3'-untranslated and 3'-flanking area of the α -globin genes) might be proportional to some earlier event such as the original α -globin gene reduplication or a subsequent correction. The sequence divergence between the two α -globin genes from the 5'-flanking area to the UAA terminator is 8 bases out of 773 or 1.1% (each postulated insertion or deletion event is given the value of one base change). We assume that the changes within IVS2 are (phenotypically) silent, based on the finding in recently diverged gene pairs that the rate of accumulation of silent changes in coding regions approximately equals that in intervening sequences¹⁹. The 1.1% silent divergence can be equated to divergence time. For recent evolutionary events in the globins (less than 85 Myr) Efstratiadis *et al.*¹³ have calculated a unit evolutionary period (UEP) of approximately 1.3 Myr (1% silent sequence divergence per 1.3 Myr) based on the method of Perler *et al.*¹⁹. Using this constant, the divergence time or time

since the last correction between the $\alpha 1$ - and $\alpha 2$ -globin genes in this homologous region is 1.4 Myr. Similarly, Zimmer *et al.* have calculated that the last correction of the α -globin genes occurred 0.7 Myr ago⁶, based on amino acid differences in several primate species. The divergence in the 3'-noncoding region and 3'-flanking regions is 17.2% (19/110) and 27.8% (17/61), respectively. If the denominators are revised to exclude the 50-base pair homologous region that encompasses the poly(A) addition site, the divergence rates become 29% (19/66) and 42% (17/40) and the estimates of the divergence time become 40.3 and 54.6 Myr, respectively. As duplication of the α -globin gene has been estimated to have occurred before the bird-mammal evolutionary divergence (~300 Myr ago)⁶, the α -globin genes apparently have corrected against each other in the 3'-untranslated and flanking regions at least once and probably numerous times since the original duplication. The observation that the α -globin gene duplication unit contains segments with different levels of homology is not unique as a similar situation has been demonstrated in the γ -globin gene family⁹. It is assumed that the regions with higher homology have undergone more recent genetic exchange.

Evidence cited here and by others^{4,10} supports the concept of recombination and exchange of DNA between defined portions of the two α -globin loci. Mapping of the α -globin genes in patients with α -thalassaemia^{20,21} and confirmatory observations *in vitro* have demonstrated that the α -globin genes can recombine by unequal cross-over at a variety of sites (ref. 4, and

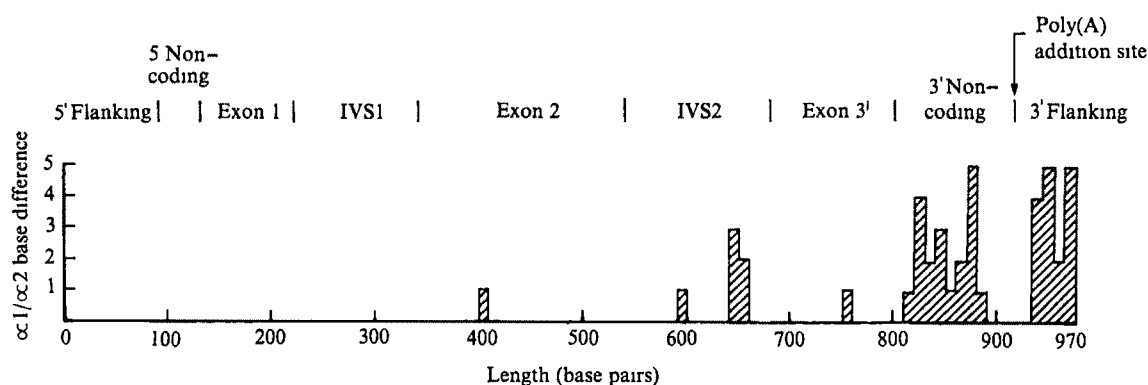


Fig. 5 Distribution of differences in the primary structure of the $\alpha 1$ - and $\alpha 2$ -globin genes. The width of each bar covers a 10-base pair segment and the height is proportional to the number of base differences within that segment. The analysis covers the α -globin genes from 88 base pairs 5' to the cap site and extends through the genes to 61 base pairs 3' to the poly (A) addition site. Each deleted base or group of bases was assigned a value of one base difference (see text).

M Goossens, unpublished data) One outcome of such a strand exchange is a chromosome with three α loci^{22,31} and a chromosome with one α locus²³. Such interactions are made possible by base pairing between two malaligned DNA strands mediated by the repeated homology units encompassing the adjacent loci. The size of these homology units may be determined by the extent of the most recent strand exchange and this in turn would determine the target size and frequency of subsequent interactions. A succession of such strand exchanges of varying lengths would yield adjacent regions that vary in degree of homology. The variation in homology observed on either side of the $\alpha 1$ - and $\alpha 2$ -globin UAA termination codons may be explained by this mechanism. Although a similar mechanism has probably maintained the homology unit common to the two γ -globin genes⁹, this sort of exchange has not occurred at a comparable rate between the closely linked β - and δ -globin genes¹³. The δ -globin gene in all primates studied is either unexpressed or expressed to a very limited extent²⁴. If limited δ -globin gene expression is due to an abnormality in its structure, then an unequal cross-over between the δ -globin and β -globin genes, such as haemoglobin Lepore, might severely diminish the β -globin gene expression leading to a β -thalassaemia phenotype. As it is deleterious, such an exchange would probably be selected against in most populations.

Our comparative sequencing of the $\alpha 1$ -globin genes from an Asian and a Caucasian shows that the sequences of these two $\alpha 1$ alleles are identical. Assuming a UEP of 1.3 Myr (see above) and a target size of 900 base pairs (the region sequenced in both $\alpha 1$ -globin genes), one base difference is expected between the α -globin alleles every 140,000 yr. (This is a minimal estimate because the UEP of 1.3 Myr is for silent changes.) From the

accumulated globin sequencing data available, it seems unlikely that the sequence of each α -globin gene has been so stringently maintained entirely by selective pressure and hence the observed suppression of polymorphism may be the result of a genetic recombination such as that postulated to occur between the γ -globin genes⁹. A mechanism for such structural correction of one allele by another (gene conversion) has been reviewed by Radding²⁵. Homologous interaction between the corresponding loci of two aligned chromosomes which contain distinct alleles can result in a non-reciprocal exchange, with the loss of one allele and its replacement by the other. Such non-reciprocal sister chromatid exchange is observed in yeast²⁶ and *Drosophila*²⁷. As racial divergence probably paralleled population segregation, genetic interaction would not be expected to occur at high frequency across racial lines. Hence the finding of an identical α -globin gene sequence in Asians and Caucasians implies that these two races probably diverged after the establishment in a common precursor population of the most recent conversion at the $\alpha 1$ -globin locus.

Mechanisms of genetic exchange cited above can be constructed to account for the observed patterns of homology within the α -globin gene region. The large repeated regions within this gene cluster probably facilitate such interaction. However, it is unclear whether any selective advantage is gained by maintenance of homology within a duplicated gene locus. How the product of these interactions becomes predominant within a population remains to be defined.

We thank Dr Nancy Cooke for helpful discussions and criticisms of the manuscript, and Kathleen Lee and Kathi A. Begley for technical assistance. This research was supported in part by grants from the March of Dimes/Birth Defects Foundation and the NIH.

Received 1 October, accepted 17 December 1980

- 1 Embury, S. H., Lebo, R. V., Dozy, A. M. & Kan, Y. W. *J. clin. Invest.* **63**, 1307-1310 (1978)
- 2 Orkin, S. *Proc. natn. Acad. Sci. USA* **75**, 5950-5954 (1978)
- 3 Deisseroth, A. *et al. Cell* **12**, 205-218 (1977)
- 4 Lauer, J., Shen, C.-K. J. & Maniatis, T. *Cell* **20**, 119-130 (1980)
- 5 Dayoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5, D56 (National Biomedical Research Foundation, Washington, DC, 1961)
- 6 Zimmer, E. A. *et al. Proc. natn. Acad. Sci. USA* **77**, 2158-2162 (1980)
- 7 Hood, L., Campbell, J. H. & Elgin, S. C. R. *A. Rev. Genet.* **9**, 305-353 (1975)
- 8 Little, P. F. R. *et al. Nature* **282**, 316-318 (1979)
- 9 Slightom, J. L., Blechl, A. E. & Smithies, O. *Cell* **21**, 627-639 (1980)
- 10 Proudfoot, N. J. & Maniatis, T. *Cell* **21**, 537-545 (1980)
- 11 Liebhauer, S. A., Goossens, M. & Kan, Y. W. *Proc. natn. Acad. Sci. USA* **77**, 7054-7058 (1980)
- 12 Jeffreys, A. J. *Cell* **18**, 1-10 (1979)
- 13 Efstratiadis, A. *et al. Cell* **21**, 653-669 (1980)
- 14 Farabaugh, P. J. & Miller, J. H. *J. molec. Biol.* **126**, 847-863 (1978)
- 15 Proudfoot, N. J., Gillam, S., Smith, M. & Longley, J. I. *Cell* **11**, 807-818 (1977)
- 16 Liebhauer, S. A. & Kan, Y. W. (submitted)
- 17 Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211-214 (1976)
- 18 Kronenberg, H. M., Roberts, B. E. & Efstratiadis, A. *Nucleic Acids Res.* **6**, 153-166 (1979)
- 19 Perler, F. *et al. Cell* **20**, 555-566 (1980)
- 20 Phillips, J. A. *et al. Blood* **54**, 1439-1445 (1979)
- 21 Embury, S. H. *et al. J. clin. Invest.* **66**, 1319-1325 (1980)
- 22 Goossens, M. *et al. Proc. natn. Acad. Sci. USA* **77**, 518-521 (1980)
- 23 Kan, Y. W. *et al. Nature* **255**, 255-256 (1975)
- 24 Martin, S. L., Zimmer, E. A., Kan, Y. W. & Wilson, A. C. *Proc. natn. Acad. Sci. USA* **77**, 3563-3566 (1980)
- 25 Radding, C. M. *A. Rev. Biochem.* **47**, 847-881 (1978)
- 26 Hurst, D. D., Fogel, S. & Mortimer, R. K. *Proc. natn. Acad. Sci. USA* **69**, 101-105 (1972)
- 27 Chunick, A. *Genetics* **75**, 123-131 (1973)
- 28 Baralle, F. E. *Cell* **12**, 1085-1095 (1977)
- 29 Proudfoot, N. J. & Longley, J. I. *Cell* **9**, 733-746 (1976)
- 30 Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. USA* **74**, 560-564 (1977)
- 31 Higgs, D. R. *et al. Nature* **284**, 632-635 (1980)

Homologous pairing can occur before DNA strand separation in general genetic recombination

Stephen C. West, Era Cassuto & Paul Howard-Flanders

Departments of Molecular Biophysics and Biochemistry, and Therapeutic Radiology, Yale University, PO Box 6666, New Haven, Connecticut 06511, USA

In the presence of ATP and Mg²⁺, purified Escherichia coli recA protein promotes the formation of joint molecules between closed circular duplex DNA and homologous circular single-stranded DNA carrying a short annealed fragment. The presence of this fragment is essential for pairing between molecules. In similar conditions recA protein is unable to act as a helicase and does not cause strand separation of the fragment from the single-stranded circle. Thus, homologous pairing between DNA molecules can take place without prior unwinding of a free end.

A GENETIC approach to the study of general genetic recombination has been to accumulate data derived from meiotic divisions in fungi and to develop models to illustrate how chromosomal DNA molecules might be cut, paired, joined and

matured into recombinant progeny¹⁻⁵. The models proposed by Holliday¹ and more recently by Meselson and Radding⁵ have been used successfully to provide a coherent explanation for extensive data on recombination in fungi^{6,7}—the early steps of

Table 1 Detection of joint molecules by the retention of form I DNA on nitrocellulose filters, after reaction with substrate DNA consisting of Φ X174 viral DNA with annealed fragment

DNA	Proteins	% Duplex DNA retained
(1) Φ X174 substrate + form I Φ X174	—	7.7
Same	recA	19
Same	recA + SSB	37
(2) Φ X174 substrate + form I plasmid	—	7.7
Same	recA	3.0
Same	recA + SSB	5.6
(3) Single-strand Φ X174 + form I Φ X174	—	6.0
Same	recA	6.0
Same	recA + SSB	9.4
(4) Single-strand Φ X174 + form I plasmid	—	10
Same	recA	5.6
Same	recA + SSB	6.5
(5) Form I Φ X174 only	recA + SSB	8.3
Form I plasmid only	recA + SSB	8.0

The assay was essentially that described by Cunningham *et al.*²⁶ Reaction mixtures of 30 μ l contained 20 mM Tris-HCl, pH 7.5, 25 mM MgCl₂, 2 mM dithiothreitol, 100 μ g ml⁻¹ bovine serum albumin, 2 mM ATP, 1.7 nmol substrate or single-stranded Φ X174 DNA, 0.2 nmol ³²P-labelled form I Φ X174 or ³H-labelled form I plasmid pBR322 DNA, 40 μ g recA protein and 2 μ g SSB protein where indicated. After incubation at 37 °C for 30 min, reactions were terminated by the addition of 2 μ l of 0.25 M EDTA and proteins removed from the DNA by the addition of SDS to 1%. Water was added to bring the volume to 200 μ l, and 50 μ l of the reaction mixtures were spotted directly onto nitrocellulose filters (Schleicher and Schuell, BA 85) to measure total counts. The remainder of the reaction mixtures was added to 3 ml of ice-cold 1.5 M NaCl, 0.15 M sodium citrate, pH 7.0, and filtered through nitrocellulose filters. These were then counted for radioactivity and the percentage of joint molecules calculated by determining the percentage of input form I DNA that was retained on the filter due to its hydrogen bonding to substrate DNA. Radioactivity due to the substrate DNA was determined in control experiments and subtracted before calculation of the percentage duplex DNA retained. In experiment 1, the difference of 29% in duplex DNA retained after incubation with recA and SSB proteins was subject to a statistical error of 3.3% (95% confidence limits). The substrate for this assay was prepared by digesting Φ X174 form I ³H-labelled DNA with *Hpa*I restriction enzyme (BRL) and separating the products on a 5% polyacrylamide slab gel. The smallest fragment (392 base pairs) was eluted from the gel, concentrated and denatured by heating, followed by rapid cooling. The fragments were then annealed with the (+) strands of Φ X174 ³²P-labelled DNA. Excess fragments not bound to the (+) strands were removed by neutral sucrose centrifugation. To measure the retention of intact circular DNA on filters that also retain the lightly labelled substrate DNA, highly labelled ³²P form I Φ X174 or ³H form I plasmid pBR322 DNA was used.

the latter model^{5,8} are shown in Fig. 1A. Recombination enzymes from fungi have been studied to a limited extent⁹, however, the cloning of bacterial recombination genes has facilitated the purification of their protein products and opened the way to an enzymological approach to the study of homologous pairing.

In *Escherichia coli*, recombination is controlled by *recA*¹⁰, the structural gene^{11,12} for a single-stranded (ss) DNA-binding protein of 38,000 molecular weight (MW)¹³⁻¹⁸ with ssDNA-dependent ATPase activity^{19,20}. Purified *recA* protein promotes homologous pairing between DNA molecules of various structures. Three situations have been studied: (1) in catalytic amounts, *recA* protein promotes hydrogen bonding between complementary single strands²¹, (2) in stoichiometric amounts, sufficient to cover all the ssDNA present, *recA* protein promotes hydrogen bonding between ssDNA fragments and homologous duplex DNA²²⁻²⁴. The product of the reaction is a D-loop, which contains a heteroduplex region with a single-strand fragment

paired to one strand of the duplex²⁵ (Fig. 1A, III). Pairing may be facilitated as *recA* protein reduces the winding number of duplex DNA²⁴, (3) *recA* protein also promotes homologous pairing between two duplex DNA molecules if one of them contains one or more short single-stranded regions^{26,27}. The latter reaction is of particular interest because genetic recombination in living cells usually involves duplex DNA molecules. Furthermore, because post-replication gaps in UV-ray-irradiated bacteria²⁸ are known to initiate sister exchanges with an efficiency approaching unity²⁹, it is likely that duplex DNA containing single-strand gaps (gapped DNA) is a natural substrate for recombination enzymes.

This important concept is supported by other results obtained with a gapped substrate. In the presence of ATP, *recA* protein recognizes gapped DNA and forms DNA-protein complexes with a high ATPase activity³⁰. If ATP is replaced by the non-hydrolysable analogue γ -S-ATP, this protein binds tightly to duplex DNA containing single-stranded regions, to form thick rod-like structures easily seen by electron microscopy. These rods are visible on gapped but not intact duplex DNA. Because homologous pairing requires ATP hydrolysis^{26,27}, we have suggested that rod-like structures formed transiently in the presence of ATP may participate in pairing between duplex molecules and the recognition of homology³⁰.

Two duplex molecules paired in a region of homology are sometimes referred to as joint molecules. The pairing of gapped DNA with intact homologous duplex DNA by *recA* protein results in joint molecules that are stable when deproteinized by SDS and proteinase K, but unstable in alkali^{26,27}. The paired molecules are linked by noncovalent bonds which show a thermal stability approaching that of duplex DNA²⁶. However, little is known about the structure in the pairing region of these joint molecules. Most published models for genetic recombination between two duplex DNA molecules assume that one duplex is

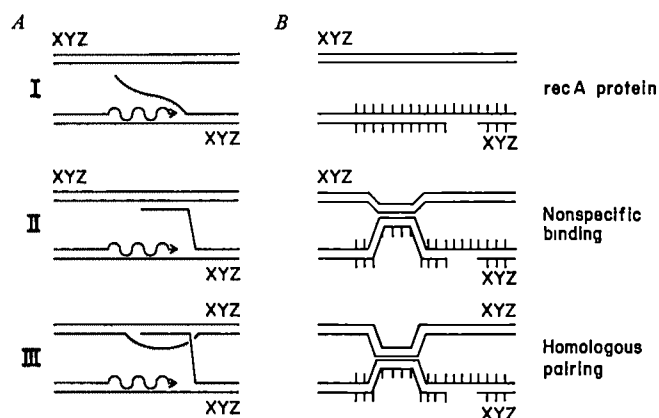


Fig. 1 Two models for the early steps in general genetic recombination. **A**, The initial steps from the model proposed by Meselson and Radding^{5,8}. I, displacement of a single strand by the replication of remaining strand. The wavy line denotes new synthesis; II, search for homology to single-stranded end; III, uptake of the free end of the displaced strand by a homologous double-stranded molecule to produce a three-stranded structure containing a heteroduplex called a D-loop^{22,25}. **B**, The initial steps of a model involving four-strand homologous pairing before strand separation. I, *recA* protein binds cooperatively to ssDNA at the site of the gap and loads the adjoining duplex; II, the loaded DNA binds nonspecifically and searches for homology in the intact duplex at the expense of ATP. The search may involve one molecule moving over the other or repeated contacts; III, formation of homologous contacts allows hydrogen bonding between the DNA molecules. The site of initial pairing may or may not include the position of the gap. The joint molecules are held by complementary hydrogen bonding, even after the removal of *recA* and SSB proteins. The subsequent steps which, in both models, include cutting, branch migration, isomerization and mismatch correction, are not shown here.

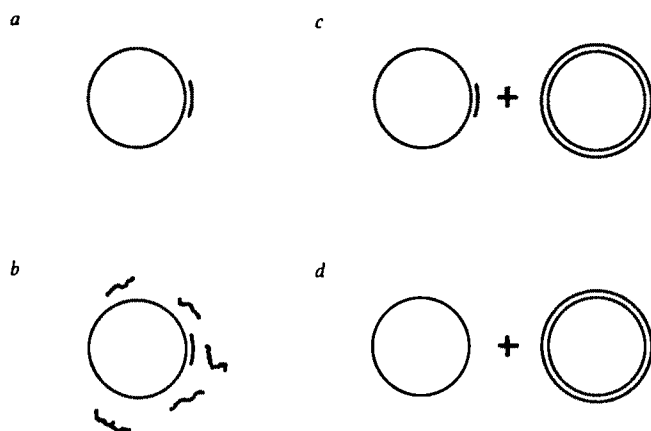


Fig. 2 DNA substrates used in the experiments *a*, Single-stranded Φ X174 viral (+) DNA with a short annealed fragment (—), *b*, the same as *a* but with addition of sonicated (+) DNA fragments, *c*, same as *a* but with intact form I Φ X174 or plasmid pBR322 duplex circular DNA, *d*, single-stranded Φ X174 DNA with intact form I Φ X174 or plasmid duplex circular DNA

cut or gapped, and the single strand with a free end is partially unwound, as shown in Fig 1A, I (ref. 5). The unwound strand is thought to act as a template in the search for complementary sequences in the second duplex. If homologous contacts are made, the free end may wind into the intact molecule to form a heteroduplex, the remaining strand of the second duplex being displaced to form a D-loop (Fig 1A, III).

However, the results of several experiments cannot be easily accounted for by a model for pairing between two duplexes which requires the initial unwinding of a free end: (1) the introduction of interstrand cross-links at both ends of the single-stranded region impairs strand separation but has no effect on the yield of joint molecules²⁷, (2) electron microscopic observations of joint molecules indicate that the joint is often distant from the single-stranded gap²⁶, (3) in addition, the experiments described here show that joint molecules can be formed by recA protein although helicase activity (unwinding and strand separation) could not be detected.

To account for these observations, we present here a working hypothesis for the initial step in homologous pairing between duplex molecules which is an alternative to D-loop formation. Unlike the Meselson–Radding model, it does not require new DNA synthesis and could therefore function in post-replication repair where replication is blocked by damaged bases^{28,29}.

RecA protein does not act as a helicase

Several proteins capable of catalysing the unwinding of duplex molecules have been characterized: helicases I, II, III and rep protein promote the unwinding of partially double-stranded (ds) DNA in ATP-dependent reactions^{31–35}. Incubation of these enzymes with a single-stranded substrate containing a short DNA duplex resulted in the separation of the strands of duplex DNA. This required the presence of single-stranded binding protein (SSB) which bound to the separated single strands and reduced the rate of renaturation³⁶. In addition, exonuclease V unwinds linear duplex molecules in the presence of SSB protein³⁷.

The substrate used in the assay comprised ³²P-labelled complementary strand fragments annealed to ³H-labelled Φ X174 viral (+) strands (Fig 2a). Figure 3 demonstrates that helicase I in the presence of ATP and SSB protein resulted in strand separation, *a* represents the sedimentation at neutral pH of the starting substrate. Although the ³²P-labelled DNA was only 194 nucleotides long, it sedimented rapidly because it was

annealed to the intact Φ X174 DNA. Figure 3b shows the results of incubation of 2 nmol DNA with 1.65 μ g helicase I for 30 min. The reaction mixture was then applied to the top of a neutral sucrose gradient and centrifuged. As can be seen, a portion of the annealed ³²P-DNA (46%) sedimented more slowly, indicating that strand separation had occurred. When SSB protein was omitted (Fig 3c), only a small amount of strand separation (<10%) was observed. Omission of ATP or helicase I gave results identical to those in Fig 3c, that is, no strand separation was observed. Very similar results were obtained when helicase I was replaced by helicase II (data not shown). These results are consistent with published data for helicase-catalysed strand separation^{31,32}.

In parallel experiments using identical conditions, the efficacy of recA protein in the strand separation assay was investigated. Figure 3d shows the results obtained after incubation of 2 nmol of substrate DNA (Fig 2a) with 20 μ g recA protein for 30 min, in the presence of SSB protein, as described above. The ³²P-labelled DNA remained annealed to the intact Φ X174 DNA which indicated that recA protein did not separate the strands. Figure 3e shows the results of a similar reaction from which SSB protein had been omitted. Reactions were also done using recA protein in the range 2–200 μ g, with or without SSB protein or ATP, or where ATP was replaced with 0.5 mM γ -S-ATP, but no strand separation was detected in any of these conditions (data not shown). We also investigated the possibility that homologous ssDNA might stimulate strand separation by recA protein. ³H-labelled Φ X174 (+) ssDNA was sonicated briefly to produce DNA fragments of average length 400 nucleotides, as estimated by sedimentation. Figure 3f shows the results of incubation of 2 nmol of substrate DNA and 4 nmol homologous (+) fragments (Fig 2b) with 40 μ g recA protein and ATP for 30 min as described. Similar results were obtained when the reaction mixture was supplemented with SSB protein (data not shown). No separation of the ³²P-labelled fragment from intact Φ X174 circles by recA protein was observed in conditions in which the separation of 10% of the fragments would have been easily detected.

To determine whether helicase-catalysed unwinding could be inhibited by the presence of recA protein, we examined its effect on the helicase II reaction in conditions of limiting helicase enzyme. Figure 4a represents the sedimentation of the starting substrate and *b* the results of incubating 2 nmol of substrate DNA (Fig 2a) with 2 μ g SSB protein and 2.4 μ g helicase II for 30 min at 37 °C. The reaction mixture was then sedimented in neutral sucrose. At this concentration of helicase II, 32% of the annealed DNA was separated from the substrate DNA and sedimented more slowly. Addition of 20 μ g recA protein to an identical reaction mixture caused a slight increase in the extent of separation (from 32 to 48%, Fig 4c), indicating that no inhibitory factors were present.

The results shown in Figs 3 and 4 show that recA protein did not act as a helicase. In similar conditions helicases I or II separated the annealed fragment from the single-stranded circle. The addition of recA protein to helicase II did not inhibit the reaction.

Formation of joint molecules by recA protein

The DNA substrate used in these experiments comprised a single-stranded circle with an annealed single-stranded fragment (Fig 2a), and is similar to a gapped duplex. Because recA protein produced joint molecules from intact and gapped duplex molecules^{26,27}, we investigated whether the substrate with a very short duplex region (Fig 2c) also formed joint molecules in the presence of recA protein. Nitrocellulose filters were used in the assay²⁶ because they allow form I DNA to pass but retain both substrate DNA and joint molecules due to their partially single-stranded nature.

To study the formation of joint molecules we incubated 0.2 nmol ³²P-labelled form I Φ X174 DNA and 1.7 nmol sub-

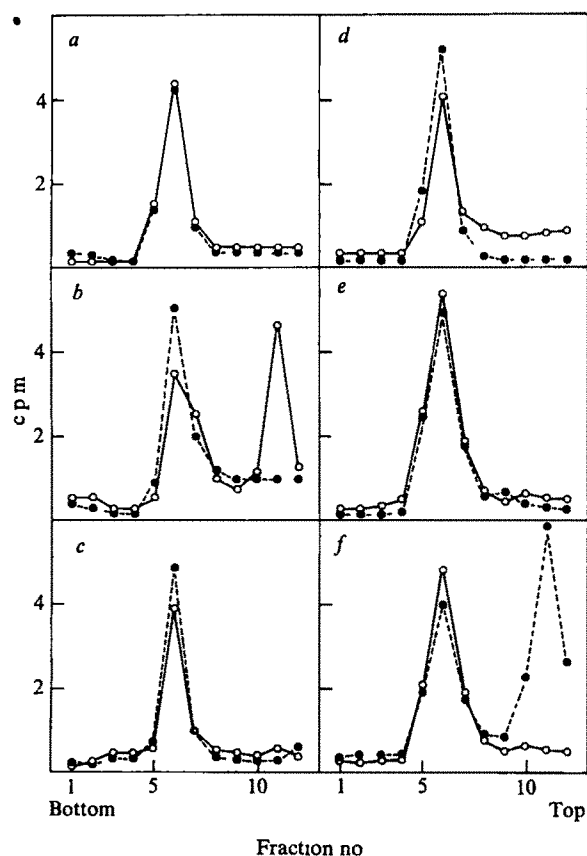


Fig. 3 Neutral sucrose sedimentation profiles showing the separation of a ^{32}P -labelled ΦX174 fragment from ^3H -labelled ΦX174 single-stranded circular DNA by helicase enzymes. The standard reaction mixture was used *a*, without enzyme, *b*, with helicase I, *c*, with helicase I but without SSB protein, *d*, with recA protein, *e*, with recA protein but without SSB protein, *f*, with recA protein and excess homologous ^3H -labelled (+) single-stranded ΦX174 fragments. Sedimentation is to the left. The substrate for the unwinding assay was prepared by digesting ΦX174 form I ^{32}P -labelled DNA with *Hae*III restriction enzyme (Bethesda Research Laboratories) and separating the products on a 5% polyacrylamide slab gel. The 194-base pair fragment was eluted from the gel, concentrated by ethanol precipitation and denatured by heating, followed by rapid cooling. The fragments were annealed with the (+) strands of ΦX174 ^3H -labelled phage DNA. Fragments not bound to the (+) strands were removed by sucrose gradient centrifugation. RecA and SSB proteins were purified as described elsewhere^{27,30}. The standard reaction mixture contained in a total volume of 100 μl , 2 nmol of substrate DNA, 20 mM Tris-HCl, pH 7.5, 25 mM MgCl_2 , 20 mM NaCl, 0.1 mM EDTA, 2 mM dithiothreitol, 2 mM ATP, 1.5 μg bovine serum albumin, 2 μg SSB protein and enzyme to be tested. Incubation was for 30 min at 37°C. The reaction was stopped by the addition of EDTA to 25 mM and SDS to 1% and, after dilution with 0.2 M NaCl to 200 μl , the mixture was sedimented through 5 ml of 5–20% neutral sucrose (10 mM Tris-HCl, pH 7.5, 1 M NaCl, 10 mM EDTA). Sedimentation was carried out in a Beckman SW 50.1 rotor at 45,000 r.p.m. and 20°C for 135 min. Fractions collected from the gradients were counted for radioactivity in Formula 963 scintillation fluid (NEN). $\circ$, ^{32}P , $\times 10^{-2}$; $\bullet$, ^3H , $\times 10^{-3}$.

strate DNA (Fig. 2c) in the presence of 40 μg of recA protein, treated the products with SDS and filtered them through nitrocellulose. The filters retained ~20% of the duplex DNA (Table 1, experiment 1). In control reactions done in the absence of any proteins or with SSB protein alone, a background of 10% of the input duplex DNA was retained. However, in the presence of both 40 μg of recA protein and 2 μg of SSB protein, almost 40% was retained. These results show that recA protein forms joint

molecules from the annealed substrate and form I DNA in conditions in which strand separation was not detected. As in other pairing reactions promoted by recA protein, ATP and Mg^{2+} were required.

The requirement for homology between molecules was also demonstrated with the filter binding assay. Joint molecules were not detected when the annealed substrate DNA and non-homologous ^3H -labelled form I plasmid DNA were incubated with recA and SSB proteins (Table 1, experiment 2). We also investigated whether single-stranded circular ΦX174 DNA, without the annealed fragment, could pair with intact circular duplex DNA (Fig. 2d). Experiments 3 and 4 in Table 1 show that incubation of single-stranded ΦX174 DNA with either ^{32}P -labelled form I ΦX174 or ^3H -labelled form I plasmid DNA (Fig. 2d) in the presence of recA and SSB proteins did not lead to retention of the duplex DNA. This shows that the partially single-stranded and partially duplex character of the substrate is needed for the formation of joint molecules.

What is the nature of the joint between the two molecules? A joint molecule could be formed by the transfer of one end of the annealed fragment to the duplex. The mechanism responsible for such a reaction might sometimes transfer the entire short fragment and so produce a D-loop which would be detected as a joint molecule in the filter binding assay. To investigate whether the observed binding of the form I DNA was due to the transfer of the ^3H -labelled fragment to the ^{32}P -labelled duplex (to form a D-loop), the reaction was stopped at the end of incubation by the addition of 1 mM $\gamma\text{-S-ATP}$ and the form I DNA was cleaved into small fragments with *Hae*III restriction enzyme. After protein removal by treatment with proteinase K and SDS, the products were sedimented through neutral sucrose. The ^3H -labelled fragments sedimented with the single-stranded circular DNA and none (<3%) was detected at the position of the ^{32}P -labelled restriction fragments which sedimented more

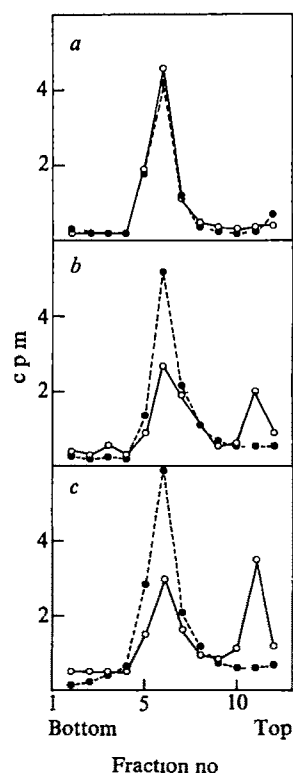


Fig. 4 RecA protein does not inhibit helicase II-catalysed strand separation. The strand separation assay was the same as that described in Fig. 3 legend. The standard reaction mixture was used *a*, without enzyme, *b*, with helicase II, *c*, with helicase II and recA protein. $\circ$, ^{32}P , $\times 10^{-2}$; $\bullet$, ^3H , $\times 10^{-3}$.

his experiment indicates that no detectable transfer of the labeled fragments to the form I DNA occurred, and will be described in greater detail elsewhere³⁸

Results indicate that recA and SSB proteins promote the formation of joint molecules between intact duplex DNA and homologous single-stranded molecules that contain a short duplication (Fig 2c). This pairing reaction occurred in conditions similar to those used in the assay for helicase activity

Homologous pairing between gapped and intact duplexes

Because most models for genetic recombination^{1-5,8} assume that the first steps in homologous pairing involve strand unwinding and the formation of a heteroduplex between the unwound strand and the complementary strand of the homologous duplex (Fig 1a), we looked for evidence that recA protein could promote strand separation

Using le-stranded circular DNA with an annealed fragment (Fig 1), we were unable to detect any helicase activity of recA protein with or without the presence of SSB protein (Fig 3). Furthermore, homologous pairing does not seem to require extensive strand separation because treatment of gapped molecules with oralen and light to produce a minimum of 1 cross-link per 50 base pairs did not inhibit the formation of joint molecules. However, we found that recA protein, which is known to promote homologous pairing between gapped circular and form I DNA^{26,27}, was also able to induce homologous pairing between form I DNA and ssDNA carrying an annealed fragment (Fig 2c, Table 1), the yield being increased by SSB protein. Although the fragment was only 392 nucleotides long, it was essential for pairing.

In this reaction, the joint molecules could be formed through the transfer of the end of the fragment to the form I duplex while the other end remained bound to the single-stranded circle, thus linking the molecules as in Fig 1A. However, it is difficult to see why the transfer of a short fragment should stop before completion. It is more likely that the transfer would go to completion and the fragment be completely assimilated into the duplex, but no such transfer was detected. If the reaction was not driven to completion, the intermediate structure would presumably dissociate through branch migration³⁹

Although we did not detect strand transfer from the annealed substrate when form I DNA was added, we have found in other reactions that recA protein promotes a strand exchange reaction if the form I DNA is replaced by blunt-ended duplex fragments

that are homologous to and overlap the ends of the annealed fragment³⁸. This exchange might take place by a reaction similar to that described for the formation of D-loops from single-stranded circular and linear duplex DNA^{38,40}

How are these observations to be reconciled? We suggest that recA protein can process some DNA substrates via D-loop formation involving strand transfer, while it may process others by four-stranded complementary pairing (Fig 1B), perhaps similar to that proposed from studies with DNA models^{41,42}, or by another mechanism involving three-stranded side-by-side pairing. Complementary hydrogen bonding may occur between two duplexes without strand separation and involves a four-stranded helical structure with contacts in the major groove⁴². We attribute the formation of joint molecules between duplex and gapped circles, described here and in recent papers^{26,27}, to these latter mechanisms

How does recA protein promote homologous pairing between two duplex molecules? The only structure known to cause exchanges with a high efficiency in intact cells is a duplex containing a post-replication gap²⁹. Purified recA protein binds cooperatively to single-stranded but not duplex DNA. It also binds to gapped duplex DNA and shows high ATPase activity, almost as great as when bound to ssDNA³⁰. To explain this we suggest that it binds initially at the gap and spreads along the adjacent duplex (Fig 1B, I), forming a bound state which may be transient as ATP is being hydrolysed. Homologous pairing can take place in conditions in which the DNA is loaded with recA protein. The initial nonspecific contact between gapped and intact molecules (Fig 1B, II) presumably leads to a search for homology and the formation of joint molecules with the base sequences aligned (see XYZ in Fig. 1B, III). Once homologous contacts are made, the DNA molecules are held together by hydrogen bonds which seem to be stable, even when deproteinized with SDS and proteinase K but are, however, alkali labile^{26,27}

The joint molecules formed by homologous pairing may then be acted on by the cutting in *trans* enzymes^{43,44}, opening the way for strand exchanges^{38,40} and formation of recombinants. Future work will determine the validity of this proposal for homologous pairing, the details of the structures involved in the joint molecules, and the way in which they are cut and joined to produce recombinants

We thank Drs H. Hoffmann-Berling and M. Abdel-Monem for helicases I and II and Drs C. DasGupta, T. Shibata, R. Cunningham and C. Radding for sending us their manuscript before publication. This work was supported by USPHS NIH grants GM 11014, CA 06519 and AM 09397

Received 2 September, accepted 8 December 1980

- 1 Holliday, R. *Genet Res* 5, 282-304 (1964)
- 2 Meselson, M. S. in *Heritage from Mendel* (ed Brink, A.) 81-104 (University of Wisconsin Press, Madison, 1967)
- 3 Hotchkiss, R. D. & A. Rev. *Microbiol* 28, 445-468 (1974)
- 4 Stahl, F. W. *Genetics* 61, 1-13 (1969)
- 5 Meselson, M. S. & Radding, C. M. *Proc natn Acad Sci USA* 72, 358-361 (1975)
- 6 Fincham, J. R. S., Day, P. R. & Radford, A. in *Fungal Genetics* 4th edn, 219-253 (University of California Press, Berkeley, 1979)
- 7 Fogel, S., Mortimer, R. K., Lusnak, K. & Taveres, F. *Cold Spring Harb Symp quant Biol* 43, 1325-1341 (1979)
- 8 Radding, C. M. *A Rev Biochem* 47, 847-880 (1978)
- 9 Holloman, W. K. & Holliday, R. *J biol Chem* 248, 8107-8113 (1973)
- 10 Willetts, N., Clark, A. J. & Low, K. B. *J Bact* 97, 244-249 (1969)
- 11 Horii, T., Ogawa, T. & Ogawa, H. *Proc natn Acad Sci USA* 77, 313-317 (1980)
- 12 Sancar, A., Stachelak, C., Kongsberg, W. & Rupp, W. D. *Proc natn Acad Sci USA* 77, 2611-2615 (1980)
- 13 McEntee, K., Hesse, J. E. & Epstein, W. *Proc natn Acad Sci USA* 73, 3979-3983 (1976)
- 14 Gudas, L. J. & Pardee, A. *J molec Biol* 101, 459-477 (1976)
- 15 McEntee, K. *Proc natn Acad Sci USA* 74, 5275-5279 (1977)
- 16 Emmerson, P. T. & West, S. C. *Molec gen Genet* 155, 77-85 (1977)
- 17 Gudas, L. J. & Mount, D. W. *Proc natn Acad Sci USA* 74, 5280-5284 (1977)
- 18 Little, J. W. & Kleid, D. G. *J biol Chem* 252, 6251-6252 (1977)
- 19 Ogawa, T., Wabiko, H., Tsurimoto, T., Masukata, H. & Ogawa, H. *Cold Spring Harb Symp quant Biol* 43, 909-915 (1979)
- 20 Roberts, J. W., Roberts, C. W., Craig, N. L. & Phizicky, E. M. *Cold Spring Harb Symp quant Biol* 43, 917-920 (1979)
- 21 Weinstock, G. M., McEntee, K. & Lehman, I. R. *Proc natn Acad Sci USA* 76, 126-130 (1979)
- 22 Shibata, T., DasGupta, C., Cunningham, R. P. & Radding, C. M. *Proc natn Acad Sci USA* 76, 1638-1642 (1979)
- 23 McEntee, K., Weinstock, G. M. & Lehman, I. R. *Proc natn Acad Sci USA* 76, 2615-2619 (1979)
- 24 Cunningham, R. P., Shibata, T., DasGupta, C. & Radding, C. M. *Nature* 281, 191-195 (1979)
- 25 Holloman, W. K., Wiegand, R., Hoessli, C. & Radding, C. M. *Proc natn Acad Sci USA* 72, 2394-2398 (1975)
- 26 Cunningham, R. P., DasGupta, C., Shibata, T. & Radding, C. M. *Cell* 20, 223-235 (1980)
- 27 Cassuto, E., West, S. C., Mursalm, J., Conlon, S. & Howard-Flanders, P. *Proc natn Acad Sci USA* 77, 3962-3966 (1980)
- 28 Rupp, W. D. & Howard-Flanders, P. *J molec Biol* 31, 291-304 (1968)
- 29 Rupp, W. D., Wilde, C. E., Reno, D. & Howard-Flanders, P. *J molec Biol* 61, 25-44 (1971)
- 30 West, S. C., Cassuto, E., Mursalm, J. & Howard-Flanders, P. *Proc natn Acad Sci USA* 77, 2569-2573 (1980)
- 31 Abdel-Monem, M., Durwald, H. & Hoffmann-Berling, H. *Eur J Biochem* 65, 441-449 (1976)
- 32 Abdel-Monem, M., Durwald, H. & Hoffmann-Berling, H. *Eur J Biochem* 79, 39-45 (1977)
- 33 Yarranton, G. T. & Gifter, M. L. *Proc natn Acad Sci USA* 76, 1658-1662 (1979)
- 34 Yarranton, G. T., Das, R. H. & Gifter, M. L. *J biol Chem* 254, 12002-12006 (1979)
- 35 Scott, J. F. & Kornberg, A. *J biol Chem* 253, 3292-3297 (1978)
- 36 Eisenberg, S., Griffith, J. & Kornberg, A. *Proc natn Acad Sci USA* 73, 3198-3202 (1977)
- 37 Rosamond, J., Telander, K. M. & Linn, S. *J biol Chem* 251, 8646-8650 (1979)
- 38 West, S. C., Cassuto, E. & Howard-Flanders, P. *Proc natn Acad Sci USA* (in the press)
- 39 Radding, C. M., Beattie, K. L., Holloman, W. K. & Wiegand, R. C. *J molec Biol* 116, 825-839 (1977)
- 40 DasGupta, C., Shibata, T., Cunningham, R. P. & Radding, C. M. *Cell* 22, 437-446 (1980)
- 41 McGavin, S. J. *J molec Biol* 55, 293-298 (1971)
- 42 Wilson, J. H. *Proc natn Acad Sci USA* 76, 3641-3645 (1979)
- 43 Ross, P. & Howard-Flanders, P. *J molec Biol* 117, 137-158 (1977)
- 44 Cassuto, E., Mursalm, J. & Howard-Flanders, P. *Proc natn Acad Sci USA* 75, 620-624 (1979)

LETTERS

H₂ emission in the EUV spectrum of T Tauri and Burnham's nebulaA. Brown*, C. Jordan*, T. J. Millar*,
P. Gondhalekar† & R. Wilson‡* Department of Theoretical Physics, University of Oxford,
1 Keble Road, Oxford OX1 3NP, UK

† Rutherford-Appleton Laboratory, Chilton, Oxon OX11 0QX, UK

‡ Department of Physics and Astronomy, University College
London, Gower Street, London WC1E 6BT, UK

Emission in the Lyman bands of H₂ has been detected in an extreme ultraviolet (EUV) spectrum of T Tauri and its adjacent nebula, obtained with the international UV explorer (IUE) satellite. We report here that the emission is in a progression which can be excited through fluorescence with the red wing of the H Ly α line, as found also in the solar atmosphere¹. The observed EUV fluxes, absence of other progressions and the flux in the previously observed² 1-0 S(1) IR line support a model where collisional rather than radiative excitation leads to an excitation temperature of $\geq 2,000$ K.

T Tau is a pre-main-sequence star, with a photospheric absorption spectrum consistent with a spectral type of K0 or K1 according to Kuhl³. It is embedded in nebulosity, from which characteristic low-excitation lines are observed in the visible region of the spectrum⁴. Emission from the $v'' = 1-0$ S(1) line of molecular hydrogen at $2.12 \mu\text{m}$ has also been detected², originating from a region ~ 5 arc s in extent centred on the star. The distance³ to T Tau is ~ 160 pc.

We observed T Tau on 4 November 1980 with the IUE satellite, from the ESA IUE ground station at Villafranca, Madrid. A short wavelength region (1,200–2,000 Å) spectrum was obtained at low spectral resolution ($\Delta\lambda \sim 6$ Å) with an exposure time of 372 min, using the large (21×9.5 arc s) aperture. The spectrum shows emission from both the star and nearby nebulosity which extends to the south. A long wavelength region (2,000–3,200 Å) spectrum was obtained at high spectral resolution ($\Delta\lambda \sim 0.2$ Å) with an exposure time of 176 min, also using the large aperture. Only the stellar emission lines of Mg II (multiplet 1) and Fe II (multiplet 1) are well exposed, many weak emission features are just detectable.

Figure 1 shows tracings through the short-wavelength spectrum of the star and nebulosity. Figure 2 shows the profiles of the stellar Mg II lines. The short wavelength spectra have been extracted from the GPHOT image by procedures discussed elsewhere⁵, and making use of the Rutherford-Appleton Laboratory IBM 360/195 computer. The long-wavelength spectrum is as provided by the IUE VILSPA processing system.

The strongest features in the short-wavelength spectrum may be attributed to transitions in O I, C II, C IV, Si II, Si III and C III

commonly observed in the EUV spectra of late-type stars² of these have been observed previously in a shorter exposure of T Tau⁶, and also in the T Tauri-type stars R U Lupi⁷ and Aur⁸.

In addition to these lines, emission features of lower intensity are observed at the wavelengths listed in Table 1. These are only lines which can be positively detected in the spectrum of the nebulosity. We identify the lines listed in Table 1 as transitions in the Lyman bands of H₂. These particular lines have been observed previously in the EUV spectrum of a sunspot⁹ it would be difficult to make a positive identification of emission without this prior knowledge. The observed wavelengths correspond to those of the strongest H₂ lines seen in solar spectrum, namely the P(5) and R(3) transitions in the $v'' = 1$ to $v' = 3, 6, 7$ and 8 bands. The laboratory wavelengths in Table 1 are taken from ref. 9. The 1-4 and 1-5 bands should be weaker, according to the calculated transition probabilities¹⁰ of the Sun.

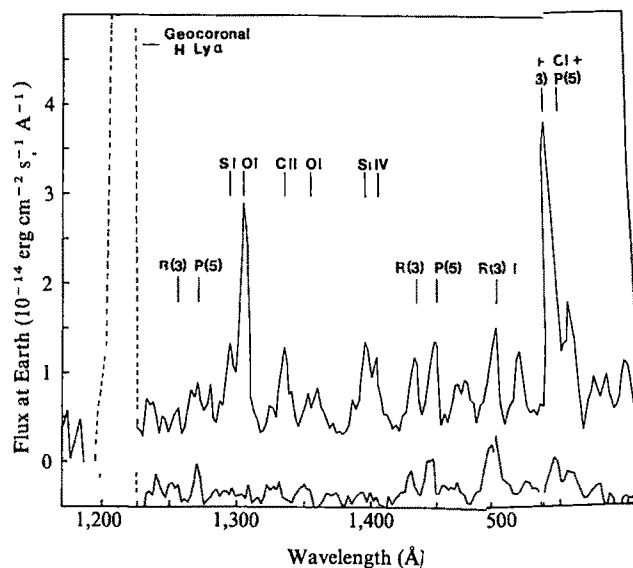


Fig. 1 Low resolution spectrum of T Tau in the region 1,200–1,600 Å. In addition to the common species indicated, the P(5) and R(3) lines of H₂ in the $v' = 1$ to $v'' = 3, 6, 7$ and 8 are observed. The upper spectrum originates in the star plus nebula in the line of sight. The lower spectrum originates in the adjacent nebulosity, and is plotted on a scale lower by $0.5 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Å}^{-1}$.

the lines are due to fluorescent emission from the upper level $v' = 1, J' = 4$, which is excited through the $v'' = 2, J'' = 5$ to $v' = 1, J' = 4$ transition, which lies $\sim 26 \text{ cm}^{-1}$ to the red of the rest wavelength (1,215.67 Å) of the atomic H Ly α line. We propose that the same excitation process is occurring in the region around T Tau. The listed fluxes include the line of sight plus the adjacent nebulosity, the latter contributing about 30% of the total. (It is not possible to distinguish between stellar and nebular emission in the line of sight.)

The presence of this emission in the absence of a strong UV continuum has important implications for the excitation of H₂ molecule in such regions. For the observed fluorescent route to be effective, a substantial population of H₂ must reside in the excited $v'' = 2, J'' = 5$ level. It has been proposed that shock heating in addition to excitation by UV radiation is required to account for the intensities of lines of H₂, both around T Tau^{11,12} and in the Orion nebula^{13–15}. Detection of the S(1) 2.12- μm line alone³ did not allow shock excitation to be distinguished from excitation by a strong EUV continuum^{16–18}. The new EUV observations favour a collisional excitation model, probably associated with shock heating. The likely relevance of H Ly α

Table 1 Lines of H₂ observed in T Tau

λ (Å)	Flux at Earth ($\text{erg cm}^{-2} \text{ s}^{-1}$)	v'	v''	Proposed identification	λ lab ⁹ (Å)
1,258	$\approx 4.2 \times 10^{-14}$	1	3	R(3)	1,257.80
1,272	8.6×10^{-14}			P(5)	1,271.93
1,431	1.0×10^{-13}	1	6	R(3)	1,431.01
1,446	1.4×10^{-13}			P(5)	1,446.13
1,490	1.8×10^{-13}	1	7	R(3)	1,489.57
1,505	2.3×10^{-13}			P(5)	1,504.79
1,547	3.6×10^{-13}	1	8	R(3) + C IV	1,547.35
1,562	1.7×10^{-13}			P(5) + C I	1,562.41

fluorescence in regions where excited H_2 exists has already been pointed out^{19,20}.

The $H\text{ Ly}\alpha$ emission visible in the IUE spectrum is due mainly to geocoronal emission, and high resolution spectra will be required to investigate the stellar component. However, the $Mg\text{ II}$ lines shown in Fig. 2 give some idea of the shape of the line profile. These lines show strong emission, with the centre of this emission shifted by 50 km s^{-1} to the red of the strong interstellar absorption feature which in this direction, should itself be redshifted by $\sim 10\text{--}20\text{ km s}^{-1}$. As the star has a radial velocity²¹ of only 22 km s^{-1} , and the width of the $Mg\text{ II}$ emission at the star will be at least the value of 120 km s^{-1} observed at the Earth, it can be argued that substantial flux should exist in the red wing of the $H\text{ Ly}\alpha$ line, in particular at $1,216.05\text{ \AA}$, the wavelength at which fluorescent excitation takes place. An estimate of the $H\text{ Ly}\alpha$ flux in the relevant wavelength range can be made by scaling from the $Mg\text{ II}$ flux, which is typically a factor of ~ 4.5 larger than $H\text{ Ly}\alpha$ in a variety of late-type stars²⁵. This gives a flux of $\sim 10^6\text{ erg cm}^{-2}\text{ s}^{-1}$ in the equivalent part of the $H\text{ Ly}\alpha$ line.

Provided all transitions from $v'=1, J'=4$ are optically thin, the measured flux in the $P(5)$ line at $1,505\text{ \AA}$ can be used in conjunction with the expression for the emissivity given by Jordan *et al.*¹⁹ to give

$$\int N(2, 5) I_\nu dV \sim 3.0 \times 10^{33} \quad (1)$$

where $N(2, 5)$ is the population density (cm^{-3}) in the level from which excitation takes place, I_ν is the intensity of the $H\text{ Ly}\alpha$ radiation field ($\text{erg cm}^{-2}\text{ s}^{-1}\text{ Hz}^{-1}$) at the nebula. Allowing for a geometric dilution factor in a medium optically thin to $H\text{ Ly}\alpha$, I_ν can be related to I_ν^* , the intensity at the star. Then

$$I_\nu \sim 1.0 \times 10^{17}/R^2 \quad (2)$$

where R (cm) is the distance at which absorption by H_2 takes place.

The observed relative intensities of the transitions to $v''=3$ and 6 or 7 suggest that the lines to $v''=3$ are marginally optically thick. However, even if the transitions to $v'' \leq 3$ are optically thick, a probability of escape argument for lines from a common upper level shows that the enhancement of the optically thin transitions is only a factor of 1.7. In view of the larger uncertainties this correction has not been applied.

The IR observations show that the H_2 emission originates from a region around the star of $\sim 5\text{ arc s}$ ($1.2 \times 10^{16}\text{ cm}$) in total extent. The spatial distribution in the optical region and in our EUV spectrum is asymmetric and less than 10 arc s in extent. We adopt an emitting volume $F4\pi R^2 dl$, where F is some fraction of the spherical shell of radius R . The column density in the 2, 5 level is then

$$\int N(2, 5) dl \sim \frac{1}{F} 2.4 \times 10^{15} \text{ cm}^{-2} \quad (3)$$

Similarly the observed flux in the $S(1)$ transition and transition probability²² can be used to give

$$R^2 \int N(1, 3) dl \sim \frac{1}{F} 2.2 \times 10^{47} \quad (4)$$

Equations (3) and (4) then give $N(2, 5)/N(1, 3)$ as a function of R , assuming that the emitting volumes are the same. The value

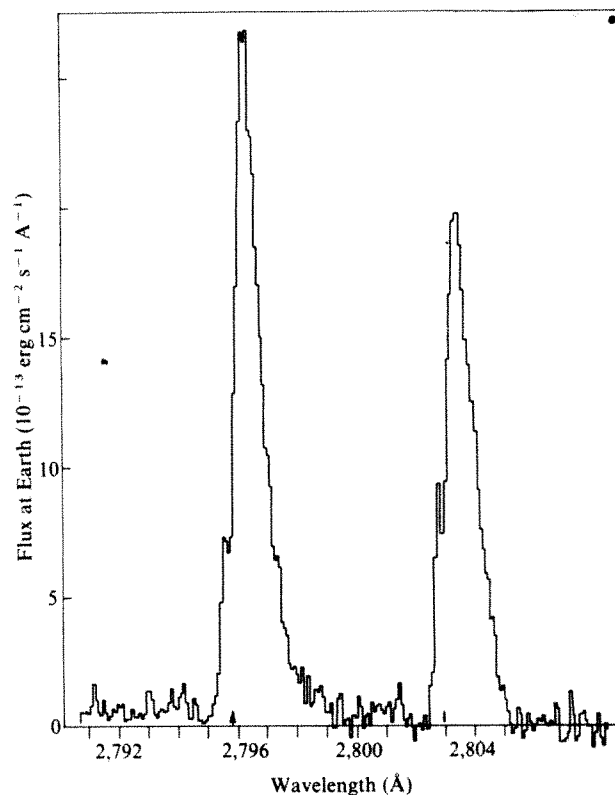


Fig. 2 High resolution spectrum of the $Mg\text{ II}$ resonance lines in T Tau. The arrows mark the rest wavelength at the star. Narrow interstellar lines appear in broad asymmetric stellar lines.

of the excitation temperature, T_{exc} , required to give this population ratio may then be found.

Table 2 gives values of R , the radius of the shell, for temperatures between 1,000 and 3,000 K. These are somewhat smaller than the value of $6 \times 10^{15}\text{ cm}$ found from the IR observations. A larger flux in the stellar $H\text{ Ly}\alpha$ emission would result in a larger value of R . Even allowing for an order of magnitude underestimation of I_ν^* would lead to $T_{\text{exc}} \geq 1,500\text{ K}$.

A further estimate of T_{exc} can be made by using the observed upper limit to the flux at $1,365.7\text{ \AA}$, where the strongest member of the $P(1)$ progression, which can be excited by $H\text{ Ly}\beta$, should occur²³. The ratio of the flux at $1,365.7\text{ \AA}$ to that in the $1,505\text{ \AA}$ line gives

$$N(2, 5)/N(0, 1) \geq 2 \times I(\text{Ly}\beta)/I(\text{Ly}\alpha).$$

Using $I(\text{Ly}\beta)/I(\text{Ly}\alpha) \sim 10^{-2}$, as in the Sun, results in $T_{\text{exc}} \geq 2,700\text{ K}$.

The mass column density (from equation (3) with $F \leq 1$), emitting mass and lower limits to $N(H_2)$ for a filled shell, are also given in Table 2. These mass column densities are consistent with lines to $v'' \leq 3$ being optically thick. Solutions with $T_{\text{exc}} \leq 1,000\text{ K}$ are excluded by the upper limit imposed by known⁴ visual extinction to the star and nebula.

Finally we may compare the results of Beckwith *et al.*² who adopted Kwan's model¹⁵ for excitation by a low velocity shock and subsequent emission through the $1\text{--}0\text{ S}(1)$ line. With $v_{\text{shock}} \sim 15\text{ km s}^{-1}$ they find $N(H_2) \sim 2 \times 10^4\text{ cm}^{-3}$. Adopting recent calculations of dissociation speeds²⁴, the upper limits to the shock velocity can be found as a function of $N(H_2)$ and are given in Table 2. In the absence of a strong EUV continuum it would be difficult to destroy the H_2 once formed.

We conclude that the EUV observations are consistent with a model where H_2 is at an excitation temperature of $\geq 2,000\text{ K}$, and that the basic excitation mechanism is collisional rather than radiative.

Table 2 Parameters of the emitting region as a function of T_{exc}

T_{exc} (K)	R (10^{15} cm)	$\int N(H_2) dl$ (cm^{-2})	$N(H_2)$ (cm^{-3})	$\mathcal{M}$ ($M_\odot$)	v_{shock} (km s^{-1})
1,000	0.38	$\geq 2.0 \times 10^{21}$	$\geq 5.3 \times 10^6$	5.6×10^{-6}	<21
1,500	0.91	$\geq 5.7 \times 10^{19}$	$\geq 6.3 \times 10^4$	9.9×10^{-7}	<22
2,000	2.10	$\geq 4.9 \times 10^{18}$	$\geq 2.3 \times 10^3$	4.4×10^{-7}	<32
3,000	3.60	$\geq 7.9 \times 10^{17}$	$\geq 2.2 \times 10^2$	2.0×10^{-7}	<51

Received 1 December 1980; accepted 9 January 1981.

- Jordan, C., Brueckner, G. E., Bartoe, J.-D. F., Sandlin, G. D. & VanHoosier, M. E. *Nature* **270**, 326–327 (1977).
- Beckwith, S., Gatley, I., Matthews, K. & Neugebauer, G. *Astrophys. J. Lett.* **223**, L 41–L43 (1978).
- Kuhi, L. V. *Astrophys. J.* **140**, 1409–1433 (1964).
- Schwartz, R. D. *Astrophys. J.* **191**, 419–432 (1974).
- Brown, A. & Jordan, C. *Mon. Not. R. astr. Soc.* (in the press).
- Gondhalekar, P. M., Penston, M. V. & Wilson, R. *Proc. Symp. The First Year of IUE* (ed. Willis, A. J.) 109–116 (University College London, 1979).
- Gahm, G. F., Fredga, K., Liseau, R. & Dravins, D. *Astr. Astrophys.* **73**, L4–L6 (1979).
- Imhoff, C. L. & Giampapa, M. S. *Astrophys. J. Lett.* **239**, L 115–L119 (1980).
- Herzberg, G. & Howe, L. L. *Can. J. Phys.* **37**, 636–659 (1959).
- Allison, A. C. & Dalgarno, A. *Atomic Data* **1**, 289–304 (1970).
- Schwartz, R. D. *Astrophys. J.* **195**, 631–642 (1975).
- Maran, S. P., Hobbs, R. W., Brown, R. L., Jura, M. & Knapp, G. R. *Astr. J.* **84**, 1709–1712 (1979).
- Hollenbach, D. J. & Shull, J. M. *Astrophys. J.* **216**, 419–426 (1977).
- London, R., McCray, R. & Chu, S. I. *Astrophys. J.* **217**, 442–447 (1977).
- Kwan, J. *Astrophys. J.* **216**, 713–723 (1977).
- Osterbrock, D. E. *Astrophys. J.* **136**, 359–364 (1962).
- Black, J. H. & Dalgarno, A. *Astrophys. J.* **203**, 132–142 (1976).
- Shull, J. M. *Astrophys. J.* **219**, 877–885 (1978).
- Jordan, C., Brueckner, G. E., Bartoe, J.-D. F., Sandlin, G. D. & VanHoosier, M. E. *Astrophys. J.* **226**, 687–697 (1978).
- Shull, J. M. *Astrophys. J.* **224**, 841–847 (1978).
- Herbig, G. H. *Astrophys. J.* **214**, 747–758 (1977).
- Turner, J., Kirby-Docken, K. & Dalgarno, A. *Astrophys. J. Suppl.* **35**, 281–292 (1977).
- Feldman, P. D. & Fastie, W. G. *Astrophys. J. Lett.* **185**, L101–L104 (1973).
- Hollenbach, D. J. & McKee, C. F. *Astrophys. J. Lett.* **241**, L47–L50 (1980).
- McClintock, W., Henry, R. C., Moos, H. W. & Linsky, J. L. *Astrophys. J.* **202**, 733–740 (1975).

A water vapour maser in the Large Magellanic Cloud

E. Scalise Jr

CRAAM/INPE: Instituto de Pesquisas Espaciais, Conselho Nacional de Desenvolvimento, Científico e Tecnológico-CNPq, CP 515, 12200-São José dos Campos, SP, Brasil

M. A. Braz

CNPq-Observatório Nacional, Rua Pará, 277, 01243-São Paulo, SP, Brasil

Surveys of the Large Magellanic Cloud (LMC), the closest satellite galaxy to our Solar System, for water vapour sources of measurable intensity have so far yielded negative results. These searches, by Johnston *et al.*¹ and Kaufmann *et al.*², with detection limits of about 150 and 20 Jy (1 Jy = 10^{-26} W m⁻² Hz⁻¹), respectively, suggest that any water vapour masers that exist in the LMC are much weaker than the sources recognized in our own Galaxy and in several more distant galaxies^{3–5}. We now report the results of a survey using the 45-ft Itapetinga radio telescope which shows the presence of a water vapour maser in the N159 region of the LMC.

The survey was carried out in 1979–80. Upper detection limits of about 1.2 Jy were attained with the use of long integration times and a maser amplifier front-end, the total system temperature ranging from 250 to 400 K. Spectral information was provided by a 47-channel spectrometer, 100-kHz bandwidth filter bank, giving a velocity resolution of 1.35 km s⁻¹. Two vertically polarized feed horns were used in beam-switching the on-off observing mode. Each independent observation was taken with 30 min integration. We selected observations made in good weather conditions and our data were corrected for atmospheric attenuation effects by assuming an average atmospheric optical depth of 0.2. The data were also corrected for random attenuation and elevation gain dependence.

We concentrated our search for the $J = 6_{16} - 5_{23}$ transition of the water maser line ($\nu = 22235.08$ MHz) in two of the strongest H(II) regions in the LMC⁶—N157 (ref. 7, 30 Doradus) and N159—and in two dark nebulae⁸—Hodge 47 and Hodge 52. Huggins *et al.*⁹ studied these four regions in their carbon monoxide survey.

Water vapour emission was detected in only one of these four positions, N159, and its spectrum is shown in Fig. 1. It presents two components, one 8 km s⁻¹ wide with a narrow feature of 7.2 Jy at 233.6 km s⁻¹, and the other with a feature of 3.2 Jy at 213.9 km s⁻¹.

It is interesting that the water maser we discovered in N159 coincides with the lowest colour temperature far-IR source detected by Werner *et al.*¹⁰ in the LMC—30 Doradus—which is also the strongest source. A large CO cloud, 24 pc in diameter, was found in this direction⁹. Other molecules such as H₂CO (ref. 11), OH (ref. 12) and CO (ref. 9), have been discovered at this position in the LMC, with velocities almost coincidental with the H₂O velocity.

Regarding the intensity of this maser, it would be a strong source of 8×10^4 Jy if placed at the distance of the Orion Nebula, but when compared with the W49 masers, it is much weaker in terms of total luminosity¹³.

A negative result was obtained for the 30 Doradus nebula, probably due to the fact that its core contains a large number of young OB stars¹⁴ responsible for the ionization of the molecules.

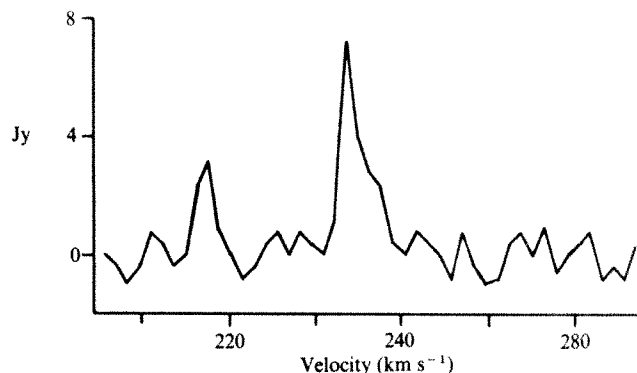


Fig. 1 Water vapour spectrum of N159 (RA 05 h 40 min 24 s DEC. $-69^{\circ}46'0.1950.0$). The main feature appears at 233.6 km s⁻¹ with respect to the Local Standard of Rest.

Hence, one should search for water farther away from its nucleus. Hodge 47 and 52 also showed no line emission.

Our search for water masers in the direction of the Large and the Small Clouds will continue, and we hope that the results improve our knowledge not only of masers but also of our closest irregular Magellanic-type galaxies.

This research resulted from a Brazil-US bilateral scientific agreement through the national agencies CNPq and NSF, respectively. The maser receiver was developed at Haystack Observatory under the direction of Dr S. H. Zisk. Partial support was received from Brazilian science agencies FINEP and FAPESP.

Received 24 November 1980; accepted 5 January 1981.

- Johnston, K. J., Knowles, S. H. & Sullivan, W. T. III *Astrophys. J.* **167**, L93 (1971).
- Kaufmann, P., Zisk, S., Scalise, E. Jr, Schaal, R. E. & Gammon, R. H. *Astr. J.* **82**, 577 (1977).
- Churchwell, E. *et al. Astr. Astrophys.* **54**, 969 (1977).
- Huchtmeier, W. K., Witzel, A., Kuhr, H., Pauliny-Toth, I. I. & Roland, J. *Astr. Astrophys.* **64**, L21 (1978).
- Marques dos Santos, P. & Lépine, J. R. D. *Nature* **278**, 34 (1979).
- McGee, R. X., Brooks, J. W. & Batchelor, R. A. *Aust. J. Phys.* **25**, 581 (1972).
- Henize, K. G. *Astrophys. J. Suppl. Ser.* **2**, 315 (1956).
- Hodge, P. W. *Publ. astr. Soc. Pacif.* **84**, 365 (1972).
- Huggins, P. J., Gillespie, A. R., Phillips, T. G., Gardner, F. & Knowles, S. *Mon. Not. R. astr. Soc.* **173**, 69P (1975).
- Werner, M. W. *et al. Mon. Not. R. astr. Soc.* **184**, 365 (1978).
- Whiteoak, J. B. & Gardner, F. F. *Mon. Not. R. astr. Soc.* **174**, 51P (1976).
- Whiteoak, J. B. & Gardner, F. F. *Mon. Not. R. astr. Soc.* **176**, 25P (1976).
- Deguchi, S. *Div. geol. planet. Sci. Contr.* 3445 (California Institute of Technology, 1980).
- Hyland, A. R., Thomas, J. A. & Robinson, G. *Astr. J.* **83**, 20 (1980).

Discrete chorus emissions observed at a low latitude ground station

Birbal Singh, Ram Prakash & Hari Singh

Department of Physics, RBS College, Bichpuri,
AGRA-283 105, India

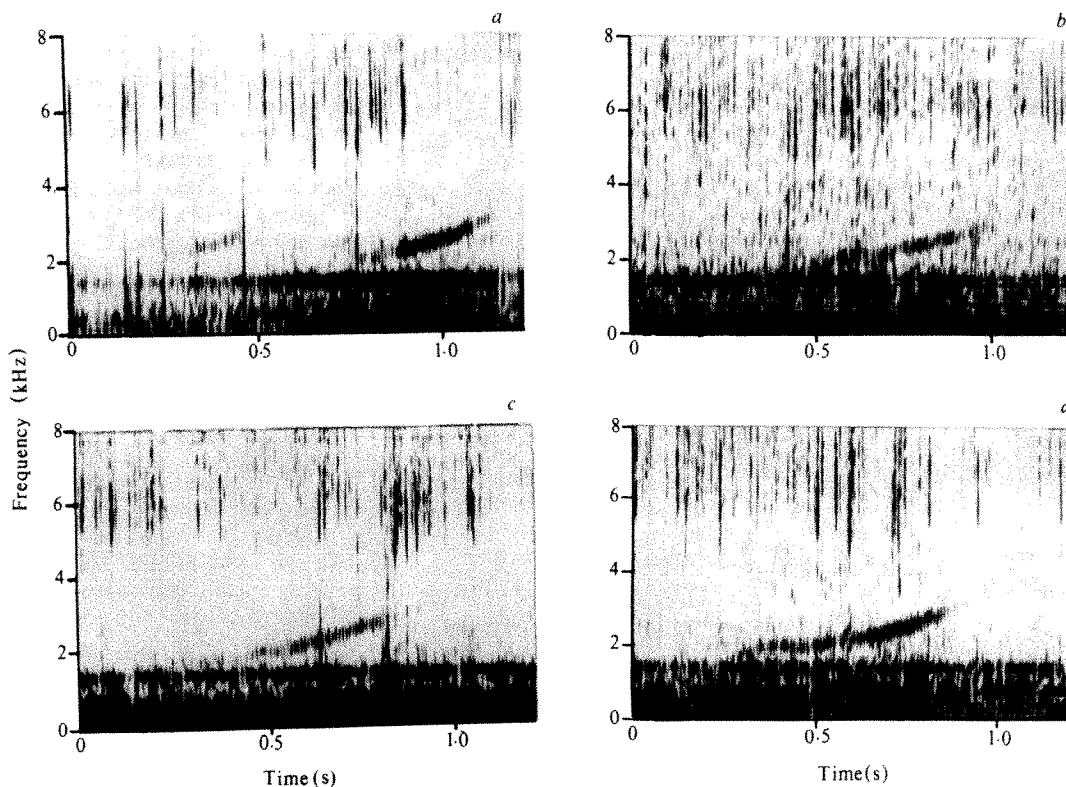
Discrete chorus emissions observed for the first time at the low latitude ground station Bichpuri, Agra (geomagnetic lat. $17^{\circ}12' \text{ N}$, $L \sim 1.1$) are reported here. The emissions are of the riser type and occurred in the narrow frequency range 1.7–3 kHz. We show that these emissions are generated in the equatorial plane near $L = 1.2$ as a result of cyclotron resonance between whistler mode waves and the inner radiation belt relativistic electrons of energy 3–5 MeV, and propagated to the low latitude ground station under the influence of the horizontal density gradients of the equatorial anomaly.

Using conventional equipment consisting of an antenna, transistorized pre and main amplifiers and a magnetic tape recorder, we started routine observations of whistlers at a low latitude ground station of Bichpuri between January 1979 and July 1979. Bichpuri is in a noise-free rural area $\sim 11 \text{ km}$ west of Agra. The observations were made at night (18.00–06.00 h LT) during the first 10 min of every hour, but, the duration of the recording was increased whenever the whistler activity was found to be high. On 12 January 1979, which was a magnetically quiet day ($K_{p\text{max}} = 3$), we observed nine discrete rising emissions between 19.25 and 20.40 h LT. Some of these emissions, shown in Fig. 1, occurred in a narrow frequency range between 1.7 and 3 kHz with the rate of change of frequency with time (df/dt) between 2 and 3 kHz s^{-1} . The 10 ms structure in the spectra of these emissions is due to the instrument. Such structures (time interval $\sim 1.5 \text{ s}$) have also been observed in high latitude periodic emissions reported by Helliwell¹ (Fig. 7.21). Figure 1a, where two rising emissions are shown, consisted of four emissions; the other two occurred somewhat later. However, all four emissions are shown on a different frequency–time record in Fig. 2. An interesting feature of these emissions is that the time interval between the consecutive emissions is approximately in a geometrical progression. Further, the frequency ranges of first

and third emissions are slightly different from those of the other two emissions.

Hitherto, discrete chorus emissions have been observed only at middle and high latitude ground stations. The lowest latitude ground station where such emissions were received is Chilbolton², Hampshire, UK (geomagnetic lat. 49.8° N , $L = 2.4$). The emissions observed by us could not have been caused by a fault in the equipment because several whistlers, with dispersions in the range $19\text{--}35 \text{ s}^{1/2}$ were also observed by the same equipment. As these riser emissions differ markedly in frequency and df/dt from those of the riser whistlers observed earlier at the low latitude ground station of Gulmarg³ (geomagnetic lat. $24^{\circ}10' \text{ N}$), the possibility that these emissions are also riser whistlers is ruled out. Further, the possibility that these emissions are high latitude discrete chorus emissions which have propagated to our low latitude ground station in the Earth-ionosphere wave-guide mode of propagation³ is also ruled out because there is a strong absorption band in the frequency range 2–5 kHz. The possibility that these emissions are generated at high L -values in the vicinity of the plasmapause and propagated to our ground station after successive magnetospheric reflections in a manner similar to those of the ELF hiss observed by satellites in the inner zone^{4,5} does not seem to be tenable. This is because, at the frequency of these emissions, the attenuation losses during various magnetospheric reflections would be high and the waves could not be detected on the ground. Although the exact losses have not been calculated, a rough estimate can be made from the work of Kimura⁶. He has shown from ray tracing computations in a model ionosphere that the amount of attenuation suffered by the waves of frequency 1 kHz from 300 km at 30° N reaching $L = 3$ after 12 successive reflections is $\sim 6 \text{ dB}$. If we include $\sim 4 \text{ dB}$ of losses suffered in the lower region of the ionosphere, then the total loss will be $\sim 10 \text{ dB}$. To reach the plasmapause near $L = 4$, the waves have to undergo some more reflections and consequently suffer more attenuation than 10 dB. As the attenuation increases with frequency, the waves of the chorus emissions generated near the plasmapause would suffer much higher attenuation than 10 dB and may not be observed on the ground. Further, at the base of the F-region ionosphere, the wave normal angles of these waves are such that the downward waves are unlikely to penetrate the lower ionosphere and reach the ground.

Fig. 1 Discrete chorus emissions observed at the low latitude ground station of Bichpuri. a, 19.27 IST; b, 19.54 IST; c, 19.48 IST; d, 19.53 IST.



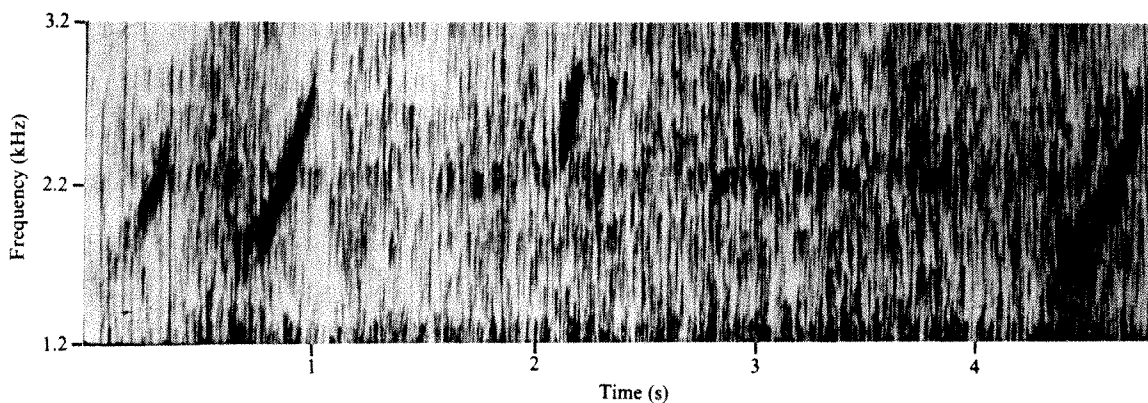


Fig. 2 Risers with time interval increasing in geometrical progression.

Thus we believe that these emissions are generated in the equatorial plane in the inner zone radiation belt ($L \sim 1.2$). Cyclotron resonance between whistler mode waves and the inner zone radiation belt electrons seems to be the possible generation mechanism for two reasons: (1) this mechanism has been found to account for the frequency-time spectra of middle and high latitude discrete chorus emissions adequately²; and (2) it also occurs in the inner radiation belt⁷. Below $L = 1.2$ electron losses are dominated by atmospheric scattering^{7,8} resulting in the small life time and density of the energetic electrons and hence the emissions observed by us cannot be generated at the top of the field line corresponding to Bichpuri. These emissions cannot be generated at L -values > 1.2 for the reasons to be stated later. To test the cyclotron resonance as the possible generation mechanism for these emissions, we calculated the resonant energy of the involved electrons and growth rates at $L = 1.2$ in the equatorial plane. The resonant energies for various frequencies of the emissions were calculated from the expression⁵.

$$E_{11} = (\gamma_r - 1)m_0c^2 \quad (1)$$

in which m_0 is the rest mass of electron, c , the velocity of light in vacuum and γ_r , the relativistic factor to be obtained from the relation⁵

$$\gamma_r - 1 \equiv \left(\frac{\tilde{\Omega}}{\omega_p}\right)^2 \left(\frac{\tilde{\Omega}}{\omega}\right) \left(1 + \frac{\Omega^+}{\omega}\right) \quad (2)$$

where $\tilde{\Omega}$ is the electron gyrofrequency, Ω^+ , the proton gyrofrequency, ω_p , the plasma frequency and ω , the wave frequency. The plasma frequency ω_p was calculated from the ionospheric model of Singh⁹ which yielded an electron density of 8.13×10^3 electrons cm^{-3} at $L = 1.2$. The calculated resonant energies, of the order of 3–5 MeV, are shown in Fig. 3a. For calculating the growth rate we followed expression (2.20) of Kennel and Pet-

schek¹⁰, which, for the case $\omega \ll \tilde{\Omega}$, reduces to

$$\gamma = \pi \tilde{\Omega} \eta A \quad (3)$$

where η is the ratio of the density of energetic electrons to that of thermal electrons and A , the pitch angle anisotropy. To obtain the density of energetic electrons at $L = 1.2$ we considered observed intensity versus L -value curves (energy > 1 MeV) given by Katz¹¹ for the inner zone radiation belt. The curve indicated a flux of 3×10^6 electrons $\text{cm}^{-2} \text{sr}^{-1} \text{s}^{-1}$ at $L = 1.2$ which yielded a density of 3–5 MeV electrons $\sim 1.26 \times 10^{-3}$ electrons cm^{-3} . A was calculated from the relation¹⁰ $A = \frac{1}{2} / \log_e(\frac{A_0}{\alpha_0})$ (α_0 being the equatorial loss cone angle, mirror height = 100 km) and found to be 2.08 which is approximately the same as calculated by Ashour-Abdalla¹² at $L = 1.2$ ($P = 250$). If we assume a pitch angle distribution of the form $\sin^m \phi$ ($m = 4$, and ϕ is the pitch angle) and compare the calculated pitch angle versus intensity curve with that of Katz¹¹, the two distributions are found to be nearly the same. Thus the value of the anisotropy $A = 2.08$ at $L = 1.2$ is justified. By substituting the values of $\tilde{\Omega}$, η and A in equation (3), the growth rates for various frequencies of the emissions were calculated and found to be $\sim 3 \text{ rad s}^{-1}$. This result indicates significant wave amplification.

Now we consider how the chorus emissions were propagated to our low latitude ground station. Suitable ducts do not exist in the low latitude ionosphere and non-ducted waves are not observed on the ground. However, Bullough *et al.*¹³ have reported the observation of VLF and ELF emissions in the Ariel 3/4 satellites inside the equatorial anomaly region. The equatorial anomaly is a well-established afternoon-evening phenomenon. Thus it is possible that the low latitude chorus emissions were propagated to our ground station under the influence of the diminishing phase of the equatorial anomaly. To verify this we calculated ray paths for the frequencies of the chorus emissions from $L = 1.2$ to the base of the F-region ionosphere (120 km) in an ionospheric model representing the diminishing phase of the equatorial anomaly. The diminishing phase of the anomaly was assumed to have a constant electron density distribution within $\pm 5^\circ$ of the equator instead of a trough at the equator. This distribution was obtained from the model $N = N_{eo} \exp \alpha (\sin \theta - \beta)$ of Singh⁹ where the value of α was taken to be 11 for the negative horizontal density gradient between 22° and 5° and zero for no ionization variation between 5° and the equator. The ray paths computed for the frequencies of the observed chorus with the initial wave normals along the geomagnetic field line at the equator ($\Delta_0 = -90^\circ$) arrived within 18.5° and 19° latitudes at 120 km, with the final wave normals almost vertically downward ($\Delta_f \approx 178^\circ$) indicating that the corresponding waves would penetrate the lower ionosphere and be observed on the ground. One of the ray paths for the frequency of 3 kHz is shown in Fig. 3b. The ray paths computed from higher L -values of 1.24 and 1.47 arrived at 24° and 45° latitudes respectively with the wave normals not vertically downward. Such waves would not be observed on the ground. Thus the observed chorus emissions were generated at $L = 1.2$ and

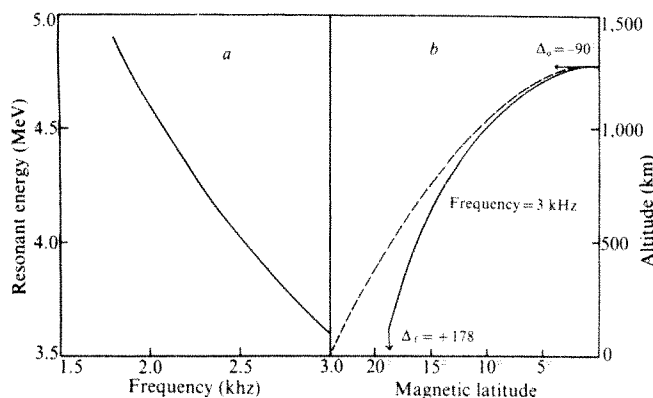


Fig. 3 a, The computed resonant energies at $L = 1.2$. b, Ray path for the frequency of 3 kHz. Δ_0 and Δ_f are the initial and final wave normal angles (angles between the wave normal and the vertical, measured from the vertical, positive in a clockwise direction).

propagated to the ground station under the influence of the favourable density gradients of the equatorial anomaly. The growth of these emissions at 1.7 kHz is due to the resonance between the energetic electrons and whistler mode waves of frequency lower than 1.7 kHz and termination of emissions at ~3 kHz is due to the absorption by the thermal background plasma in the interaction region¹⁴. The detailed mechanism in which the four emissions of Fig. 2 were generated with two of them having different upper and lower cutoff frequencies is under investigation and the result will be published elsewhere.

The occurrence of natural whistlers and emissions is a rare phenomenon at our ground station. During the 7 months of observation, we recorded whistlers on four occasions and emissions on one occasion only. The rare occurrence is believed to be due to non-availability of suitable ducts in the low latitude ionosphere. The emissions observed on 12 January 1979 between 19.25 and 20.40 h LT are due to fortuitous presence of favourable density gradients of the equatorial anomaly in the low latitude ionosphere. These emissions, although received only once at our ground station, indicate their origin in cyclotron resonance at low L -values in the ionosphere. Unfortunately, satellite data on inner zone electrons corresponding to the period during which the emissions were observed are not available to confirm the result. However, more data on these emissions and simultaneous satellite data on energetic electrons would be useful for studying the pitch angle scattering and precipitation of energetic particles as a result of cyclotron resonance at low L -shells. We have therefore continued the observations of these emissions in our ground station.

We thank Dr D. L. Carpenter for valuable suggestions and Dr S. S. Agrawal for help with the sonagraph.

Received 7 May; accepted 12 December 1980.

- Helliwell, R. A. *Whistlers and Related Ionospheric Phenomena* (Stanford University Press, 1965).
- Rycroft, M. J. *Radio Sci.* **7**, 811 (1972).
- Dikshit, S. K., Somayajulu, V. V. & Tantry, V. A. P. *Nature* **230**, 115 (1971).
- Muzzio, J. L. R. & Angerami, J. J. *J. geophys. Res.* **77**, 1157 (1972).
- Tsurutani, B. T., Smith, E. J. & Thorne, R. M. *J. geophys. Res.* **80**, 600 (1975).
- Kimura, I. *Radio Sci.* **1**, 269 (1966).
- Imhof, W. L., Gaines, E. E. & Reagan, J. B. *J. geophys. Res.* **78**, 4568 (1973).
- Imhof, W. L. & Smith, R. V. *Radiation Trapped in the Earth's Magnetic Field* (ed. McCormac, B. M.) 100 (Reidel, Dordrecht, 1966).
- Singh, B. J. *J. geophys. Res.* **81**, 2429 (1976).
- Kennel, C. F. & Petschek, H. E. *J. geophys. Res.* **71**, 11 (1966).
- Katz, L. *Radiation Trapped in the Earth's Magnetic Field* (ed. McCormac, B. M.) 130 (Reidel, Dordrecht, 1966).
- Ashour-Abdalla, M. *Magnetospheric Physics* (ed. McCormac, B. M.) 241 (Reidel, Dordrecht, 1974).
- Bullough, K., Hughes, A. R. W. & Kaiser, T. R. *Magnetospheric Physics* (ed. McCormac, B. M.) 231 (Reidel, Dordrecht, 1974).
- Liemohn, H. B. *J. geophys. Res.* **72**, 39 (1967).

High propagation rates of explosions in large volumes of gaseous mixtures

S. R. Moore & F. J. Weinberg

Department of Chemical Engineering and Chemical Technology, Imperial College, Prince Consort Road, London SW7 2BY, UK

To account for the level of damage observed following unconfined vapour cloud explosions of the kind that occurred at Flixborough in 1974, flame speeds of the order 10^2 m s^{-1} have had to be postulated^{1,2}. Detonations (in which reaction is initiated by adiabatic compression in a shock wave) are not thought to be involved, whilst deflagrations (in which molecular and turbulent transport processes from the flame bring the reactants to a state of rapid reaction) do not seem capable of producing sufficiently high velocities. We suggest here that radiation from a large enough body of burnt products could be sufficient to produce multi-point initiation ahead of the flame by igniting particles, especially those consisting of loose aggregates of fine fibres. Although we thought initially that the hazard would chiefly be due to particles of flammable materials, such as cotton waste and other 'fluffy' aggregates of cellulosic substances, we show below that non-flammable materials are also hazardous.

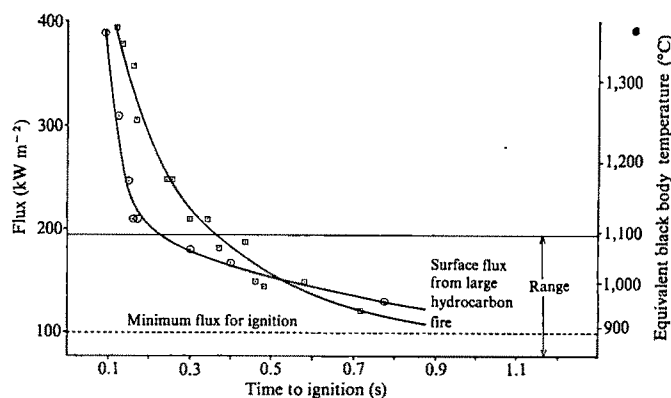


Fig. 1 Minimum ignition times for propane-air (~3.8% propane) ignition by Kaowool (O) and cotton wool particles (□).

Whilst the emissivities of combustion products are very small at optical path lengths encountered in normal practice, they become appreciable for the very large clouds of burnt and burning gases characteristic of vapour cloud explosions. The black-body radiation corresponding to stoichiometric flame temperatures is of the order of 1 MW m^{-2} . In view of the difficulties of arranging and monitoring large scale explosions of premixed reactants, existing measurements have generally been confined to large experimental fires³⁻⁵, usually burning pools of liquid fuels (see, however, ref. 5). These are diffusion flames in which burning does not occur uniformly throughout the gas but is confined to a relatively thin peripheral layer. Radiation fluxes measured near such flames are of the order of $100\text{--}200 \text{ kW m}^{-2}$.

It is easy to show that quite sporadic ignitions ahead of the main front would yield a sufficiently enhanced flame speed. Defining flame speed in terms of the rate of consumption of a given volume of reactants—which also defines the increased rate of generating volume of hot products—multiple ignitions increase total flame area and decrease the distance each flame front has to cover to complete burn-out. From simple geometrical considerations the 'magnification' of the apparent flame speed resulting from multiple ignitions, separated by distance x across a cloud of linear dimension X , is approximately $2X/x$ times the normal propagation rate, and this does not take into account additional effects of turbulence. Note that the present hypothesis is not intended to supersede other possible mechanisms which have been postulated. We wish rather to draw attention to a very significant additional effect which must be taken into account if it occurs at all.

A few order-of-magnitude calculations were carried out on the likely rates of temperature rise of particles subjected to such radiation, to check whether the concept was reasonable. Using a mass/irradiated area ratio of 0.02 kg m^{-2} and a specific heat capacity of $1.7 \text{ kJ kg}^{-1} \text{ K}^{-1}$ (typical for cellulosic materials) rates of temperature rise of the order $10^4\text{--}10^5 \text{ K s}^{-1}$ are obtained. Typical auto-ignition temperatures would, therefore, be expected to be reached in $10^{-2}\text{--}10^{-1} \text{ s}$. Moreover, there are likely to be cooperative effects between flammable mixture and heated particle as regards the ignition criterion: the ignition temperature or ignition energy may be considerably lower for a combustible particle in a flammable mixture than that for either a combustible particle in air or for an inert particle in a combustible mixture. It therefore seems reasonable that the available radiation flux could raise the temperature of a particle, irrespective of its size (that is, over dimensions larger than the quenching distance) to a reasonable ignition temperature in the time during which a sufficient radiation source persists.

The short duration of the ignition time lag means that it is the growth of the initial radiation source to a size giving sufficient emissivity that is likely to take the longest time. Thereafter ignition of all the centres irradiated at the same flux may be virtually simultaneous, irrespective of the time lag affecting each

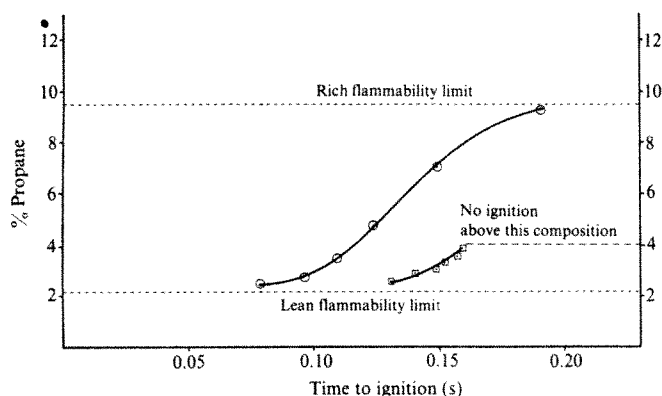


Fig. 2 Effect of gas mixture composition on minimum ignition times (flux = 350 kW m^{-2}) of Kaowool (○) and cotton wool particles (□).

of them. Due to the attenuation of radiation with distance, such initiation will, however, occur progressively ahead of the advancing flame front. The over pressures generated will depend on this rate of progression but the contribution of radiation ignition will not fall to zero until the propagation rate is so fast that the ignition lag, integrated over the range of radiation fluxes ahead of the front, remains incomplete by the time the front reaches the particle.

To study the concept, various particles were irradiated in a flammable atmosphere by a small continuous-wave CO_2 laser constructed in these laboratories with a nominal power of 20 W. The distribution of radiation flux across the beam was determined calorimetrically and peaked at 400 kW m^{-2} . The irradiated particles were suspended on quartz fibres above a burner mouth discharging pre-mixed propane-air in various proportions, metered by rotameters. The time to ignition was measured as a function of radiation flux and mixture strength for each type of particle. The onset of ignition was monitored by means of an ionization probe connected to an oscilloscope. Synchronization of particle irradiation with the start of the time measurement was achieved by placing a steel shutter in the laser beam and connecting this to the trigger input of the oscilloscope. Some results are shown in Figs 1 and 2.

As might be expected the dependence of the results on the form of aggregate of the particles is more complex than simple theory suggests. For compact aggregates such as small pieces of paper, charring occurs without producing ignition in our experimental conditions. However, loose aggregates of fine fibres, such as cotton wool, will ignite in a small fraction of a second at a radiation flux $< 10\%$ of that corresponding to the black-body radiation at stoichiometric flame temperatures (see Fig. 1) and well within the range of values measured near large fires. Note that non-flammable materials in this form of aggregate, such as the refractory 'Kaowool' produce ignition at least as readily (in fact in less time—see Fig. 2).

Presumably the effectiveness of such particles in producing ignition is due to their low heat capacity for a given volume and area irradiated. The mechanism is now being studied in greater detail. In this context, note that the fibre diameters of Kaowool and cotton wool are $5 \mu\text{m}$ and $20 \mu\text{m}$ respectively. It also seems that the evolution of blanketing vapour from the fibres must be considered in parallel with the rise of temperature in time. Thus it seems likely that the rich limit of ignitable compositions of the mixture is extended to the flammability limit in the case of Kaowool (see Fig. 2) because, unlike cotton wool, it does not evolve any flammable vapour.

The conclusion that materials in this form of aggregate are readily ignited by levels of radiation flux corresponding to those measured in large experimental fires, seems to us of far-reaching importance. Note that the kind of aggregate most conducive to ignition by radiation is also the most likely to be raised by a blast

wave or convection current, due to its low free-fall velocity. If the hypothesis is correct, it suggests that the practical hazard may be alleviated, for example by spraying and thus retaining dangerous dusts on wetted surfaces.

We thank Professor H. Gg. Wagner for introducing F.J.W. to this problem and for many helpful discussions, and Dr R. W. Dibble for building the CO_2 laser.

Received 30 October 1980; accepted 5 January 1981.

1. Briscoe, F. *Canvey Island Rep.* Appendix 1 (HMSO, London, 1978).
2. Guggan, K. *Unconfined Vapour Cloud Explosions* (Institution of Chemical Engineers, London, 1978).
3. Brown, L. E., Wesson, H. R. & Walker, J. R. *Hydrocarb. Process.* **53**, 141 (1974).
4. Nash, P. *Symp. Chem. Proc. Hazards V*, Paper D. (Institution of Engineers, 1974).
5. Hardie, H. C., Lee, D. O. & Benedick, W. B. *Combust. Sci. Technol.* **17**, 189 (1978).

Decomposing the IR absorption spectra of diamonds

Gordon Davies

Wheatstone Physics Laboratory, King's College, Strand, London WC2R 2LS, UK

Nitrogen is the most important impurity found in natural diamonds, affecting properties as diverse as colour¹ and thermal conductivity². Consequently it is important to distinguish between the different molecular forms adopted by the nitrogen. Improved data for obtaining the relative concentrations of two of the most common molecular forms are now presented. These data are then applied to a set of diamonds taken from one source and it is confirmed that the relative concentrations of the different forms of the nitrogen are intimately related to each other.

Most natural diamonds (perhaps 98%)³ contain their nitrogen predominantly in the A molecular form, believed to be a pair of nearest neighbour substitutional atoms⁴, as well as in other aggregated forms. These other forms are the N_3 and N_9 centres which may be identified by UV spectroscopy¹, the B aggregates and the platelets. The platelets produce⁵ a clearly resolved absorption peak at $1,370 \text{ cm}^{-1}$ ($7.3 \mu\text{m}$) as shown in Fig. 1. However, the A and B aggregates give overlapping absorption (Fig. 1). In the spectral region of Fig. 1, pure diamonds are essentially transparent.

To separate the A and B spectra we can take the ratio $m = \mu(7.8)/\mu(8.5)$ of the absorption coefficients measured at $7.8 \mu\text{m}$ ($1,280 \text{ cm}^{-1}$) and at the peak absorption near $8.5 \mu\text{m}$ (near $1,175 \text{ cm}^{-1}$). From m we can obtain the ratio $r = \mu_{7.8}^A/\mu_{7.8}^B$ of the absorption produced separately by the A and B aggregates at $7.8 \mu\text{m}$. Although existing data⁶ relating m and r have been used⁷⁻⁹ these were based on a very few largely unsuitable diamonds and are slightly incorrect when r is large. For the present study diamonds were selected for good homogeneity, with large faces (about $3 \times 5 \text{ mm}^2$) and with optimum absorption strengths (giving minimum transmissions of about 40% of the incident light). It is now found that

$$r = (2.72m - 1)/(1 - 0.41m) \quad \text{when } m < 1.5$$

and

$$r = (1.88m - 1)/(1 - 0.5m) \quad \text{when } m > 1.5$$

(It is preferable to use two expressions, depending on the value of m , as the wavelength of the peak absorption near $8.5 \mu\text{m}$ varies with the A to B admixture, as is obvious from Fig. 1.) Hence

$$\mu_{7.8}^A = \mu(7.8)r/(1+r); \quad \mu_{7.8}^B = \mu(7.8)/(1+r) \quad (1)$$

where $\mu(7.8)$ is the total absorption measured at $7.8 \mu\text{m}$. An alternative way to obtain r (for example in strongly absorbing diamonds) is to use the absorption ratio m' measured at $9.9 \mu\text{m}$

(1,015 cm⁻¹) and 9.1 μm (1,100 cm⁻¹), $m' = \mu(9.9)/\mu(9.1)$, giving

$$r = (0.795 - 1.164m') / (0.124m' - 0.0333)$$

and

$$\begin{aligned}\mu_{7.8}^A &= \mu(9.1)r / (0.125r + 1.164); \\ \mu_{7.8}^B &= \mu(9.1) / (0.125r + 1.164)\end{aligned}\quad (2)$$

Equations (1) and (2) are believed to give $\mu_{7.8}^A$ and $\mu_{7.8}^B$ correct to $\pm 5\%$ of the total absorption at 7.8 μm.

As an application of these results we consider the relationship obtained by Brozel *et al.*⁷ on the platelet A and B absorption in different diamonds. They worked with specimens taken from arbitrary sources. As the distribution is determined thermodynamically¹⁰ we may obtain cleaner statistics by using specimens from one growth environment. Forty-seven diamonds from an eclogitic environment¹¹ in the Premier Mine (South Africa) gave the relationship shown by the dots in Fig. 2 for the variation of platelet absorption at 7.3 μm divided by $\mu(7.8)$ with the fraction of $\mu(7.8)$ produced by the A aggregates. (All these quantities are measurable directly from the IR spectra.)

For $\mu_{7.8}^A/\mu(7.8) > 0.5$ the data lie on a remarkably uniform trend with three times less scatter than reported by Brozel *et al.*⁷. However, for smaller values of this ratio the data deviate from Brozel's curve⁷ (shown by the line in Fig. 2 after transforming to the axes used here and adjusting the ordinate to allow for the arbitrary scale used in ref. 7). Supplementing the Premier

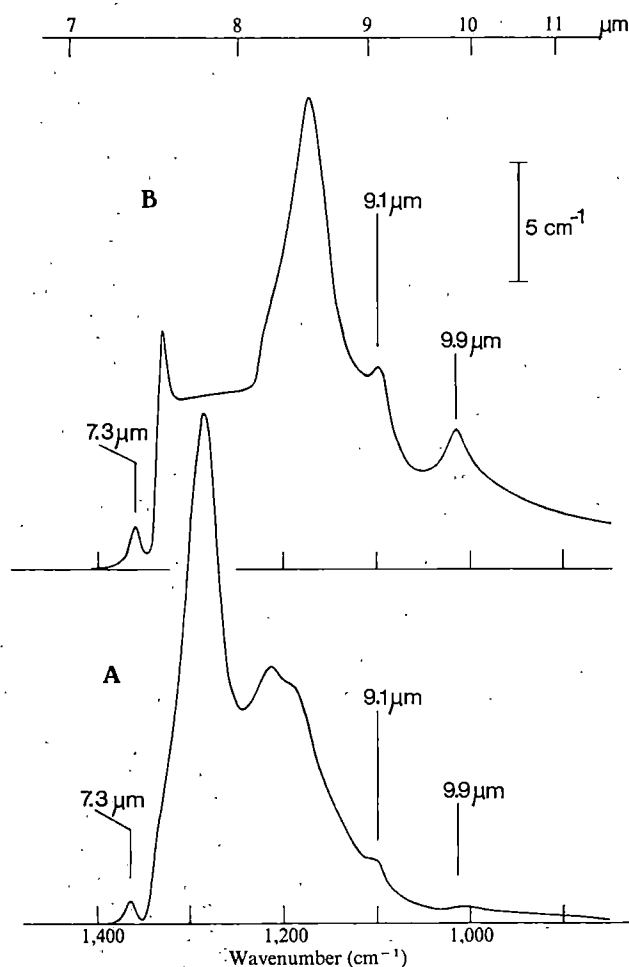


Fig. 1 IR absorption spectra of diamonds containing predominantly A and B aggregates.

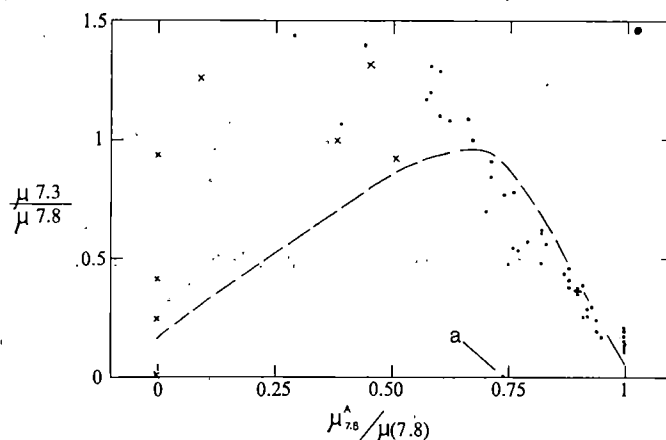


Fig. 2 The variation of (platelet peak absorption divided by total 7.8 μm absorption) with the ratio r of the absorption at 7.8 μm produced separately by the A and B aggregates. Diamond A has a total 7.8 μm absorption of $\mu(7.8) = 0.5 \text{ cm}^{-1}$, far lower than that of most of the other specimens, which had $5 \leq \mu(7.8) \leq 50 \text{ cm}^{-1}$. ●, Diamonds from the Premier Mine; ×, random species. The line is taken from ref. 7 after transformation and scaling.

diamonds with random specimens produces a scatter and not the unambiguous approach to the origin given by Brozel *et al.*⁷.

The significance of these results is not limited solely to the somewhat esoteric relationships plotted in Fig. 2. It has been shown¹²⁻¹⁴ that single substitutional nitrogen atoms are able to migrate through the diamond lattice at temperatures around 1,800 K, apparently by a vacancy-enhanced diffusion process¹⁴. The immediate end product of the migration is self-trapping to form the A aggregate of nitrogen. Likewise the A aggregates may be broken into single atoms by heat treatment (for example, at 2,250 K)⁷. The equilibrium established by the nitrogen between the single and the paired form depends on the temperatures experienced by the diamonds¹⁰. The other aggregated forms of nitrogen mentioned above also take part in the equilibration process^{9,10}. Viewed in this context, the relationships shown in Fig. 2 suggest that, in principle, the relative concentrations of nitrogen in a given diamond can be used to estimate the temperatures experienced by that diamond since its formation; the conditions prevailing at its formation can be determined from the nature of the inclusions trapped within it at the time of growth¹¹. As well as being of interest to the geologist, this approach is important to the gem industry because the restructuring of the nitrogen impurity by heat treatment can improve the colour of a diamond¹⁷.

I thank J. W. Harris for the Premier Mine diamonds, Z. E. Ciechanowska for running preliminary spectra and A. T. Collins for experimental advice.

Received 4 November 1980; accepted 16 January 1981.

1. Davies, G. *Chem. Phys. Carbon* **13**, 1-143 (1977).
2. Burgemeister, E. A. *J. Phys. C* **13**, L963-968 (1980).
3. Dyer, H. B., Raal, E. A., du Preez, I. & Loubser, J. H. N. *Phil. Mag.* **11**, 763-774 (1965).
4. Davies, G. *J. Phys. C* **9**, L537-542 (1976).
5. Sobolev, E. V., Lisoivan, V. I. & Lenskaya, S. V. *Soviet Phys. Crystallogr.* **12**, 665-668 (1968).
6. Davies, G. & Summersgill, I. *Diamond Research*, 6-16 (Industrial Diamond Information Bureau, London 1973).
7. Brozel, M. R., Evans, T. & Stephenson, R. F. *Proc. R. Soc. A* **361**, 109-127 (1978).
8. Burgemeister, E. A. *Physica* **93B**, 165-179 (1978).
9. Davies, G. *J. Phys. C* **5**, 2534-2542 (1972).
10. Evans, T. *Diamond Research* 17-22 (de Beers, Ascot, 1978).
11. Harris, J. W. *Diamond Research* 2-6 (de Beers, Ascot, 1979).
12. Chrenko, R. M., Tuft, R. E. & Strong, H. M. *Nature* **270**, 141-144 (1977).
13. Collins, A. T. *J. Phys. C* **11**, L417-422 (1978).
14. Collins, A. T. *J. Phys. C* **13**, 2641-2650 (1980).
15. Evans, T. & Rianey, P. *Proc. R. Soc. A* **334**, 111-130 (1975).
16. Allen, B. P. thesis, Univ. Reading (1980).
17. General Electric Company, UK Patent 1578987 (filed 30 June 1977).

Quantitative retention of magmatic argon in a glassy basalt

David E. Fisher

University of Miami, Rosenstiel School of Marine and Atmospheric Science, 4600 Rickenbacker Causeway, Miami, Florida 33149, USA

Basalts erupted in deep waters commonly chill sufficiently rapidly to form a glass outer skin. Due to the rapid chilling and the high hydrostatic pressures, magma-ambient gases and in particular the rare gases may be trapped within this glass¹⁻³. Several theoretical investigations have examined the relationship of the trapped $^{40}\text{Ar}/^{36}\text{Ar}$ ratio to the history of terrestrial degassing and formation of the atmosphere⁴⁻¹¹. These measurements have not yet, however, answered the fundamental questions of whether the mantle $^{40}\text{Ar}/^{36}\text{Ar}$ ratio varies significantly either vertically within the Earth and/or from one geographic locality to another, or even whether a quantitative estimate of the magmatic concentration and isotopic ratio at any given location is in fact obtainable. This report deals with the last of these questions. Our data indicate that the answer is in the affirmative.

In any sample, the Ar actually measured will consist of unknown proportions of a trapped magmatic and a contaminating atmospheric component, and it is possible that the trapped magmatic component may not be quantitatively retained in the

Table 1 Ar data measured in seamount P-50 (5°59'S, 98°09'W)

Sample no.	Description	^{36}Ar ($10^{-10} \text{ cm}^3 \text{ g}^{-1}$)	$^{40}\text{Ar}/^{36}\text{Ar}$ ($10^{-10} \text{ cm}^3 \text{ g}^{-1}$)
1	0.68 g glass; mildly heated	0.3 ± 0.1	275
2	Above sample melted	2.05	8,370
3	1+2	2.35	7,366
4	0.4 g glass; mildly heated	0.99	295
5	Above sample melted	1.54	11,300
6	4+5	2.53	6,960
7	1.2 g glass; mildly heated	1.1	455
8	Above sample melted	1.85	8,890
9	7+8	2.95	5,760
10	Interior rock	4.3	1,570
11	Interior rock	8.4	1,880

glass. In what follows, $^{40}\text{Ar}_a$ and $^{40}\text{Ar}_m$ denote atmospheric and the unknown mantle ^{40}Ar , while m and a denote their ratios ($^{40}\text{Ar}_m/^{36}\text{Ar}_m$) and ($^{40}\text{Ar}_a/^{36}\text{Ar}_a$) respectively, the former being, of course, unknown. Then¹²

$$\begin{aligned} \left(\frac{^{40}\text{Ar}}{^{36}\text{Ar}} \right)_{\text{measured}} &= \frac{^{40}\text{Ar}_m + ^{40}\text{Ar}_a}{^{36}\text{Ar}_{\text{measured}}} \\ &= \frac{m^{36}\text{Ar}_m + a(^{36}\text{Ar}_{\text{measured}} - ^{36}\text{Ar}_m)}{^{36}\text{Ar}_{\text{measured}}} \\ &= \frac{(m-a)^{36}\text{Ar}_m}{^{36}\text{Ar}_{\text{measured}}} + a \end{aligned} \quad (1)$$

If, therefore, the measured $^{40}\text{Ar}/^{36}\text{Ar}$ plotted against $(^{36}\text{Ar})^{-1}$ gives a straight line for a suite of samples, $(m-a)^{36}\text{Ar}_m$ must be constant for those samples and the concentration and/or isotopic ratio of the trapped component may then be obtainable. Of two early sets of data, one^{1,13,14} gave two reasonably straight lines while the other¹⁵ showed only random scatter. A later set of data from a DSDP core gave an excellent straight-line fit¹⁶. These data are summarized in Fig. 1. Recently Dymond *et al.*¹⁷ combined these data with their own unpublished data and concluded that they show merely random scatter below the upper-limit envelope (Fig. 1a); such behaviour would indicate post-eruptive diffusive loss from the glassy basalts and limit the retrievable information to only a lower limit of the product $(m-a)^{36}\text{Ar}_m$.

It is important to determine whether samples exist which have quantitatively trapped the magmatic Ar without such subsequent diffusive loss, so that reliable estimates of the concentration and isotopic ratio may be determinable in principle and so that the search for vertical and/or lateral differences may be profitably pursued. These data on several determinations of Ar concentrations from one dredge haul, seamount P-50 from the south-east Pacific, (previously described¹) fit the straight-line relationship of equation (1) and thus indicate quantitative retention of the trapped Ar. It is important that this contention refers to the magmatic gases present in the lava as it erupts onto the surface of the Earth. Processes of elemental fractionation¹⁸ or vesiculation¹⁹ preceding eruption might enhance or deplete the elemental concentrations in the glass phase, but would not affect the isotopic ratios.

The data of Table 1 were obtained using a stainless-steel Reynolds-type rare gas mass spectrometer by a process similar to that described in ref. 20. The samples were chips taken from the glass crust of a single basalt flow. Blank levels were $\sim 2 \times 10^{-11} \text{ cm}^3$ for ^{36}Ar ; accuracy, $\pm 10\%$; precision of the $^{40}\text{Ar}/^{36}\text{Ar}$ ratio, $\sim 2\%$. Samples 1, 4 and 7 were heated for 30 s, sufficient to raise the crucible to a red glow, in an attempt to separate the (surface-adsorbed) atmospheric from the internally trapped components. The technique was at least partially successful:

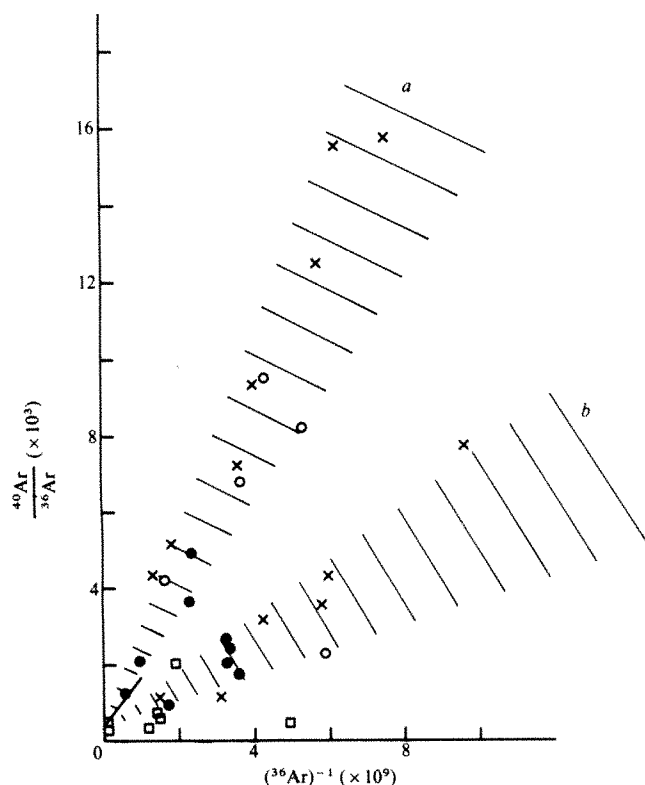


Fig. 1 $(^{40}\text{Ar}/^{36}\text{Ar})_{\text{measured}}$ plotted against $(^{36}\text{Ar})^{-1}$. The data of refs 1, 13 and 14 were measured on glasses from the East Pacific Rise and the Vema Fracture Zone. The solid line shows the slope and extent of samples from a DSDP Cretaceous core taken from the southern end of the Bermuda Rise. ●, Ref. 1; ○, ref. 13; × ref. 14; □, ref. 15.

runs 1 and 4 show only atmospheric Ar, while run 7 contains a small amount of magmatic Ar, identified by the higher $^{40}\text{Ar}/^{36}\text{Ar}$ ratio. None of the samples 1, 4 and 7 would be expected to fall on the 'constant-magmatic' line, as they obviously cannot reflect a constant-magmatic Ar component; indeed, they all lie far off the curve, to the lower right of Fig. 2. As only zero or insignificant amounts of magmatic Ar were removed from the samples in these runs, both the high-temperature fractions and the total abundances should plot along the constant-magmatic line, as, in fact, they do. Samples 10, 11, 18 and 19 are interior rock specimens of P-50; such interior rock are expected not to retain quantitatively the magmatic gases²¹, and indeed these samples lie below the constant-magmatic line.

Figure 2 includes previously published data from this same seamount; the total data bank thus includes measurements on several different samples of this dredge haul, made with three different mass spectrometers at two different laboratories over 12 yr. The straight-line fit to the glass samples (shown in Fig. 2) is remarkably good over the entire range of measurements ($^{40}\text{Ar}/^{36}\text{Ar} = 1,282\text{--}11,300$). Anchored at an intercept of 295,

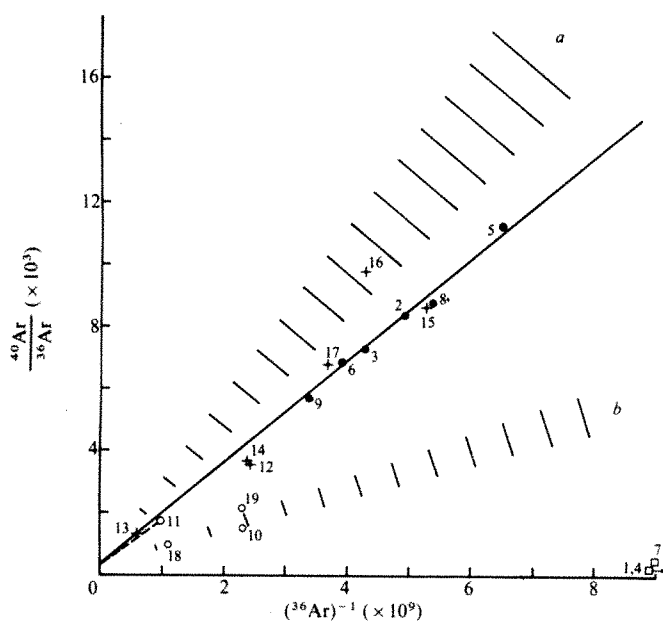


Fig. 2 The envelopes *a* and *b* are from Fig. 1. Samples 1–11 are from Table 1. Sample 12 is from ref. 14; samples 13 and 14 from ref. 1; samples 15–19 from ref. 13. ●, Glasses, this work; +, glasses, previous work; ○, interior rock; □, first (mild) heating.

corresponding to the model relationship

$$(^{40}\text{Ar}/^{36}\text{Ar})_{\text{measured}} = (m - a) \cdot ^{36}\text{Ar}_m / ^{36}\text{Ar}_{\text{measured}} + a$$

the excellent straight-line fit indicates a constant value for $(m - a) \cdot ^{36}\text{Ar}_m$. In this equation both m and $^{36}\text{Ar}_m$ are unknown. Figure 1 shows that a similar relationship was found by Takaoka and Nagao for DSDP samples from Hole 417D of Leg 51 (Cretaceous basalt flows in the Caribbean¹⁶). They solved equation (1) for m by assuming that either all the ^{36}Ar they measured was trapped magmatic, or alternatively that $^{36}\text{Ar}_m = 5 \times 10^{-10} \text{ cm}^3 \text{ g}^{-1}$ in all the samples. The first assumption provides only an upper limit to $^{36}\text{Ar}_m$; it is obviously an overstatement, as it denies the existence of the constant-magmatic line demonstrated by their data. The second assumption is based on Craig and Lupton's²¹ analysis of the early data^{14,15,22}, which include both atmospheric and magmatic Ar with no attempt to differentiate between the two; the result being no more than an average of all total measured values. This cannot possibly reflect

an accurate world-wide average of the $^{36}\text{Ar}_m$ component alone, as is emphasized by Table 1, in which many of the samples show significantly smaller concentrations of total measured ^{36}Ar . I prefer to rewrite equation (1) as:

$$(^{40}\text{Ar}/^{36}\text{Ar})_{\text{measured}} = \frac{(m - a)}{m} \cdot \frac{^{40}\text{Ar}_m}{^{36}\text{Ar}_{\text{measured}}} + a \quad (2)$$

and then recognize that the maximum measured $^{40}\text{Ar}/^{36}\text{Ar}$ datum is necessarily a lower limit to m , so that from sample 5 $m \geq 11,000$, that is $m \gg a$ and $(m - a) \sim m$. The slope of the line then gives the $^{40}\text{Ar}_m$ concentration directly: $172 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1}$. With $m \geq 11,000$, the concentration of $^{36}\text{Ar}_m \leq 1.5 \times 10^{-10} \text{ cm}^3 \text{ g}^{-1}$.

For the Leg 51 DSDP data the analysis is not as straightforward. We can say that the $^{40}\text{Ar}_m$ must be $\geq 143 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1}$; if the $^{36}\text{Ar}_m \leq 1.5 \times 10^{-10}$, then $m \geq 9,500$. On the other hand, the limit to m set by these data alone is $m \geq 1,700$, and then $^{40}\text{Ar}_m \leq 173 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1}$ and $^{36}\text{Ar}_m \leq 10 \times 10^{-10}$. The mantle component therefore looks similar, but not necessarily identical, to that recorded in the P-50 glass. The similarity in both concentration and isotopic ratio allowed by the data for Ar erupted on the other side of the world and 100 Myr previously is striking.

The data from seamount P-50 show that: (1) Samples do exist which have trapped the magma-ambient rare gas component without post-eruptive diffusive loss. (2) In the region sampled by the P-50 magma, the mantle ratio $^{40}\text{Ar}/^{36}\text{Ar}$ is $\geq 11,000$, more than an order of magnitude higher than several other estimates of the mantle ratio^{4,10,23,24}. (3) Any one measurement of the amount and isotopic ratio of Ar in a given sample is not quantitatively significant, as atmospheric contamination is generally present. Also, unless a straight-line relationship among several samples can be demonstrated, diffusive loss of the mantle component cannot be discounted.

If further measurements show this straight-line relationship in suites of samples erupted from varying depths and geographic positions, and in varying tectonic conditions, we should be able to determine whether or not a reasonably unique worldwide mantle component exists, and thus whether models of atmospheric evolution and terrestrial degassing based on such mantle-wide parameters have a chance of being valid. Alternatively we may be able to document any vertical and/or lateral variations in this component, and thus provide information pertinent to models which require such variations¹⁷.

Samples of P-50 were supplied by E. Bonatti. Research was supported in part by the NSF.

Received 13 October 1980; accepted 9 January 1981.

1. Funkhouser, J. G., Fisher, D. E. & Bonatti, E. *Earth planet. Sci. Lett.* **5**, 95–100 (1968).
2. Noble, C. S. & Naughton, J. J. *Science* **162**, 265–267 (1968).
3. Dalrymple, G. B. & Moore, J. G. *Science* **161**, 1132–1135 (1968).
4. Brown, J. F., Harper, C. T. & Odom, A. L. *Nature* **250**, 130–133 (1974).
5. Alexander, E. C. Jr & Schwartzman, D. W. *Nature* **259**, 104–108 (1976).
6. Hamano, Y. & Ozima, M. *Terr. Rare Gases* 155–171 (1978).
7. Ozima, M. *Geochim. cosmochim. Acta* **39**, 1127–1134 (1975).
8. Schwartzman, D. W. *Nature* **245**, 20–21 (1973a).
9. Schwartzman, D. W. *Geochim. cosmochim. Acta* **37**, 2479–2495 (1973).
10. Tolstikhin, I. N., Drubestkai, E. R. & Sharashin, A. Ya. *Geokhimiya* No. 4, S 14–20 (1978).
11. Fisher, D. E. *Terr. Rare Gases* 173–183 (1978).
12. Dymond, J. & Hogan, L. *Earth planet. Sci. Lett.* **38**, 117–128 (1978).
13. Fisher, D. E. *Earth planet. Sci. Lett.* **9**, 331–335 (1970).
14. Fisher, D. E. *Nature* **256**, 113–114 (1975).
15. Dymond, J. & Hogan, L. *Earth planet. Sci. Lett.* **20**, 131 (1973).
16. Takaoka, N. & Nagao, K. *Init. Rep. DSDP Legs 1–3* 1121–1126 (1980).
17. Hart, R., Dymond, J. & Hogan, L. *Nature* **278**, 156–159 (1980).
18. Fisher, D. E. *Nature* **282**, 825–827 (1979).
19. Schwartzman, D. W. *Terr. Rare Gases* 185–192 (1978).
20. Fisher, D. E. *Geochim. cosmochim. Acta* (in the press).
21. Craig, H. & Lupton, J. E. *Earth planet. Sci. Lett.* **31**, 369–385 (1976).
22. Fisher, D. E. *Geophys. Res. Lett.* **1**, 161–164 (1974).
23. Tolstikhin, I. N. *Terr. Rare Gases* 33–62 (1978).
24. Manuel, O. K. *Terr. Rare Gases* 85–91 (1978).

ESR dating of ancient flints

E. G. Garrison* & R. M. Rowlett

Department of Anthropology, University of Missouri, Columbia, Missouri 65211, USA

D. L. Cowan & L. V. Holroyd

Department of Physics, University of Missouri, Columbia, Missouri 65211, USA

Natural radioactivity can produce measurable amounts of radiation damage in rocks and glasses, and the gradual accumulation of this damage with time provides a built in 'clock' that can be used as a dating tool for geology and archaeology. Although a studied technique for estimating this damage involves measurements of stored thermoluminescent (TL) activity¹⁻⁴, several authors have reported the direct observation of natural radiation damage by ESR⁵⁻⁷. We describe here ESR measurements on a damage signal which we have found to be present without exception in a wide variety of natural flints. The point defect responsible for the signal is identified and the samples are calibrated with an artificial damage simulation. Preliminary age determinations are presented for several natural flints. Flint artefacts heated in antiquity will have their clock 'reset' to zero, and measurements on one sample suggest that the ESR technique may have applications in the study of early cultures.

Flint is a hydrothermally grown microcrystalline silica rock. Typical samples contain α quartz, amorphous silica, various silica allomorphs and a few per cent of non-silica 'binder' material. The ESR spectrum generated by natural radioactivity is shown in Fig. 1a. This spectrum is very similar to that found in quartzes artificially damaged by fast neutron bombardment⁸,

damage signal shown in Fig. 1d has two major components; the sharp structure between $g = 2.000$ and $g = 2.0020$ associated with a dangling Si bond (an E' centre), and a broad but well defined, $S = 1/2$, powder pattern at lower field. Ignoring the E', we find the low field signal grows smoothly and uniformly with fast neutron damage, and thermally anneals away as a unit. Pure γ -ray radiation (^{60}Co , up to 3 MR) has little effect on this signal in natural flints. It gives no measurable increment to the natural signal, and in samples where the natural signal has been annealed away, the weak low field signal produced by γ rays has a distinctly different appearance. (An interesting exception is when the natural signal is only partially annealed away, when even a small γ -ray dose will reconstitute a fraction of the original signal before saturating. Presumably, structural precursors are being repopulated here.) An examination of the g values of the natural signal suggests an interpretation for this insensitivity to γ rays. The values agree well with those of the recently identified 'peroxy centre'¹¹, an O_2^- radical bonded at one end to a silicon atom. If we accept this identification, then the defect is an interstitial oxygen, and a displacive lattice collision is needed to produce it.

Whether the structural damage is due to α -recoils or fast neutrons, the ultimate source of point defects is a cascade of displacement collisions in the wake of an energetic (kiloelectron volts) oxygen or silicon ion (a primary knock-on or PKO), and to first order the only parameter required for calibration is the energy deposited in the lattice. For the natural damage this is a straightforward computation. Each emission leads to a recoil nucleus with some 80 keV of energy that ultimately finds its way to the lattice through PKOs. Summing over all decays (^{238}U , ^{232}Th and their daughters in secular equilibrium) leads to an energy deposition rate

$$P_r = 8.12 \times 10^5 (\text{MeV yr}^{-1} \text{ cm}^{-3}) \times C_{\text{eff}} \quad (1)$$

where C_{eff} is the effective radioactive impurity concentration in $\mu\text{g per g}$, given by $C_{\text{eff}} = C_{\text{U}} + 1/3 C_{\text{Th}}$.

Table 1 ESR ages of some geological flints

Sample	Original flint	Accepted geological age (Myr)	t (s)	T	T	P	C_{eff} ($\mu\text{g per g}$)	ESR age (Myr) $\pm 30\%$
				C_{U} ($\mu\text{g per g}$)	C_{Th} ($\mu\text{g per g}$)	C_{U} ($\mu\text{g per g}$)		
3 g	Oklahoma	370	1,507	2.25	3.07	0.26	0.67	310
5 g	Utah	250	1,854	3.00	0.415	0.41	0.70	366
7 g	Africa	3,200(?)	304	0.693	0.802	0.52	0.74	57
9 g	Missouri	320	700	1.11	0.209	0.18	0.29	333
Ill	Illinois	320	1,347	1.20	1.20	0.42	0.664	280

Column 4 gives the average reactor exposure needed to regenerate the native ESR signal, with all data scaled to row 1, site H, reactor conditions. The effective radioactive concentration in column 8 is computed from the measurements of total concentration and powder only concentration in the preceding three columns, as described in the text. The uncertainty quoted in the final column includes statistics and a rough estimate of possible systematic errors.

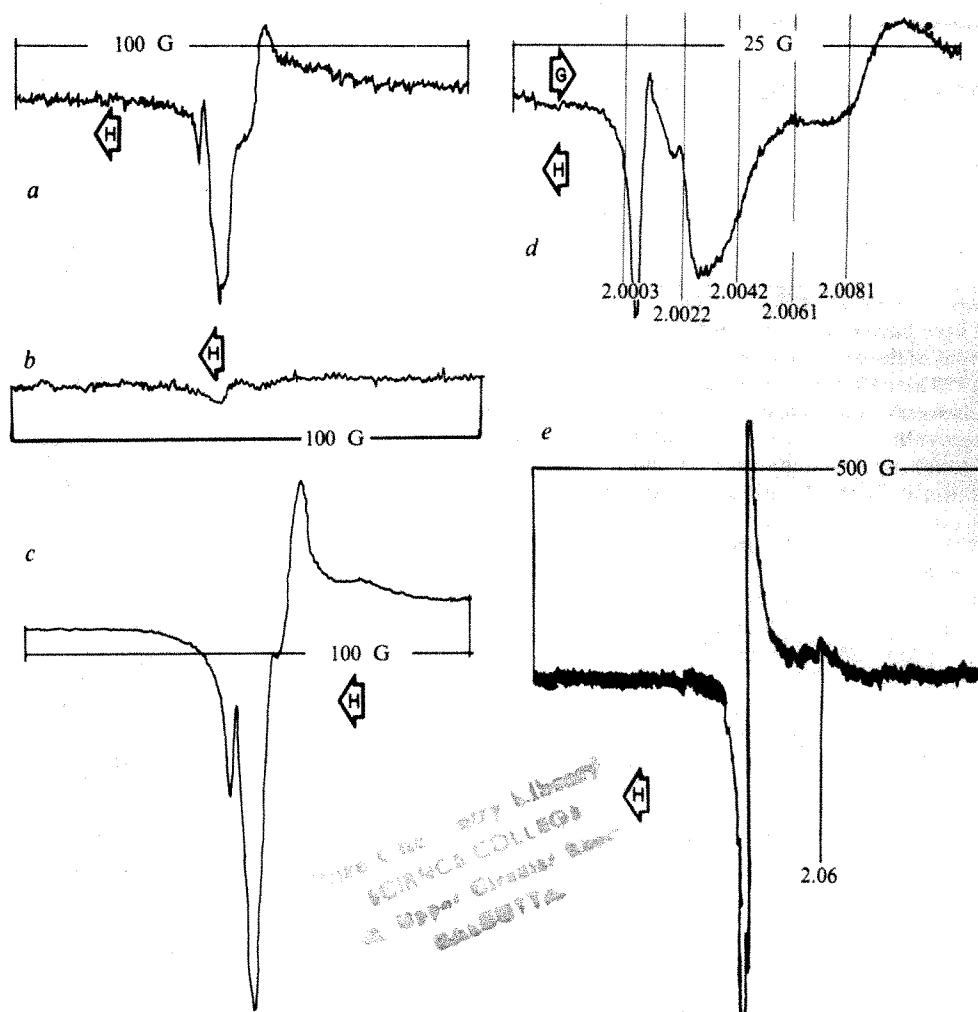
and also to the native signal observed by McMorris⁵ in igneous (but not in hydrothermal) quartz minerals. We believe, with McMorris, that this signal appears because of structural damage in the silica lattice due to the momentum-conserving nuclear recoils of α -emitting impurities. (β , γ and even α particles dissipate nearly all their energy in electronic processes and do little structural damage.) The signal thermally bleaches at 500 °C (Fig. 1b) and is readily regrown by neutron damage (Fig. 1c).

Recent work on radiation-damaged silica^{9,10} has led to a fairly clear picture of some of the simpler point defects in the silica lattice. The expanded, high-resolution trace of the natural

The total ^{238}U and ^{232}Th concentrations were determined directly by neutron activation analysis. Because flint is heterogeneous, we need to correct these measurements for the fraction of radioactivity in the binder instead of the silica lattice. We estimate this partitioning by crushing the samples and etching them in hot aqua regia. This procedure leaves a fine white powder of silica microcrystals, but removes the binder. ESR of the powder shows that etching removes impurity signals (particularly Mn^{2+}) but leaves the original peroxy signal only slightly reduced. The powder is then re-assayed for ^{238}U and we assume the less important ^{232}Th is distributed the same way. It is probably wrong to neglect the radioactivity in the binder entirely, because some of these atoms will be registered on the crystal interfaces where they can still damage the lattice (the recoil range is $\sim 200 \text{ \AA}$), so we include the binder fraction in our estimates of C_{eff} , but weighted with an efficiency of 0.1.

* Present address: Cultural Resources Laboratory, Texas A & M University, College Station, Texas 77843, USA.

Fig. 1 Absorption derivative ESR signals in Illinois flints. $\nu = 9.1$ GHz, $H_0 = 3,200$ G. *a*, The natural signal, $T = 300$ K. *b*, The same sample after annealing at 773 K. *c*, The same sample after exposure to radiation damage by fast neutrons. Correcting for gain and modulation, the signal is 12 times larger than the natural signal in *a*. *d*, An expanded high-resolution trace near $g = 2$, showing powder pattern turning points near $g_1 = 2.002$ and $g_2 = 2.008$. *e*, A high-gain, high-modulation trace at $T = 100$ K, showing the powder pattern turning point near $g_3 = 2.06$.



To calculate the energy deposited by fast neutrons, the neutron flux distribution must be evaluated at the irradiation site. We have done this by combining reactor-design calculations and measuring thermal neutron fluxes. (In the Missouri University Research Reactor, 27.8 cm from the centre of the core row 1, site H, we find $(d\phi/dE)dE = (k/E)dE$, $E < 3$ MeV, with $K = 4.5 \times 10^{12}$ neutrons per $\text{cm}^2 \text{s}^{-1}$.) These relatively low-energy collisions are nearly isotropic and the total energy given to PKOs is readily computed¹². However, the PKO spectrum has a significant high-energy component, so some of this energy is lost in irrelevant electronic processes. Evaluating this effect with a standard model^{13,14}, we find an overall efficiency of 0.36 for producing structural damage, and an energy deposit rate in the lattice (in row 1, site H of our reactor) of

$$P_n = 1.12 \times 10^{11} \text{ MeV cm}^{-3} \text{ s}^{-1} \quad (2)$$

Combining equations (1) and (2) then leads to

$$\text{Age (Myr)} = 0.139 t(s) / C_{\text{eff}} (\mu\text{g per g}) \quad (3)$$

where t is the reactor exposure time needed to reproduce the native signal.

This last expression was used in a blind study of several old flints. The results are presented in Table 1, along with the accepted geological ages. Except for the Precambrian African sample (an unusual flint subjected to obscure geological processing), the agreement in ages seems very satisfactory.

As expected, flints known to have been heated in the past (by reduced TL) showed a marked reduction in the peroxy signal. One heated sample was investigated more carefully. It came from Combe Grenal France, in an Acheulian archaeological stratum deposited $\sim 2.5 \times 10^5$ yr BP (ref. 15). The ESR spectrum near $g = 2.00$ is contaminated by impurities, but the long plateau

region extending towards $g = 2.06$ (Fig. 1e) shows clear evidence for an elevated absorption when compared with a sample heated in the laboratory. The characteristic peroxy signal reappears after neutron bombardment, and a (lengthy) extrapolation suggests that the reactor exposure time needed to reproduce the original signal would be about 5 s. The total ^{238}U concentration is $2.5 \mu\text{g per g}$, and if this U is partitioned normally between the silica and binder then the maximum elapsed time since heating would be about 5×10^5 yr. Because some of the plateau signal observed could come from centres other than the peroxy, or from an unannealed residue, the heating could well have been more recent. The main refinement needed for dating young samples is not so much an improved signal-to-noise ratio, but rather an ability to discriminate fresh peroxy signal from interfering background. This refinement may prove to be within the capabilities of standard ESR techniques.

The sample from Combe Grenal France was obtained from Professor Francois Bordes.

Received 1 August; accepted 28 November 1980.

1. Aitken, M. J. *Physics in Archaeology* 2nd edn (Clarendon, Oxford, 1974).
2. Fleming, S. *Thermoluminescence Techniques in Archaeology* (Oxford University Press, 1979).
3. Hackens, T., Aitken, M. J. & Mejdahl, V. (eds) *Pact Vol. 3* (Council of Europe, 1979).
4. Wintle, A. G. & Huntley, D. J. *Nature* **279**, 710 (1979).
5. McMorris, E. W. *Nature* **226**, 146 (1970).
6. Ikeya, M. *Nature* **255**, 48 (1975); *Archaeometry* **20**, 149 (1978); *Science* **207**, 977 (1980).
7. Robins, G. V., Seeley, N. J., McNeil, D. A. C. & Symons, M. R. C. *Nature* **276**, 703 (1978).
8. Weeks, R. A. *J. appl. Phys.* **27**, 1376 (1965).
9. Griscom, D. L. *J. non-crystalline Solids* **31**, 241 (1978).
10. Friebele, E. J. & Griscom, D. L. *Treatise Mater. Sci. Technol.* **17**, 257 (1979).
11. Friebele, E. J., Griscom, D. L., Stapelbroek, M. & Weeks, R. A. *Phys. Rev. Lett.* **42**, 1346 (1979).
12. Thompson, M. W. *Defects and Radiation Damage in Metals* 125 (Cambridge University Press, 1969).
13. Kinchin, G. H. & Pease, R. S. *Rep. Prog. Phys.* **18**, 1 (1955).
14. Thompson, M. W. *Defects and Radiation Damage in Metals* 151, 157 (Cambridge University Press, 1969).
15. Bordes, R. *A Tale of Two Caves* (Harper & Row, New York, 1972).

Stones, pits and Stonehenge

M. W. Pitts

Alexander Keiller Museum, Avebury, Marlborough, Wilts SN8 1RF, UK

There have been many mathematical or astronomical explanations of the construction of certain elements at Stonehenge¹⁻⁹. In 1979 a small salvage excavation at the site resulted in the chance discovery of a prehistoric pit dug to take a standing stone. The recognition that such substantial structures can exist unknown at Stonehenge is important to those who seek alignments or patterns in the locations of neolithic features. That 'known' arrangements could be affected by such unexpected discoveries is here illustrated with a preliminary reassessment of the significance of the Heelstone and the Station Stones (a full description of all aspects of Stonehenge is given in ref. 10).

In May-June 1979, a trench 24×1.5 m was excavated along the proposed route of a telephone exchange cable on the roadside verge (Figs 1, 2). Besides the expected segments of the circular ditch around the Heelstone and the eastern ditch of the Avenue, a completely unexpected prehistoric feature was found. This comprised part of a wide pit with the impression on its base of the stone it had once held upright. No direct dating evidence was obtained, but the pit apparently relates to early activity at the site. The standing stone was removed, and the pit backfilled and fully consolidated before the ditch around the

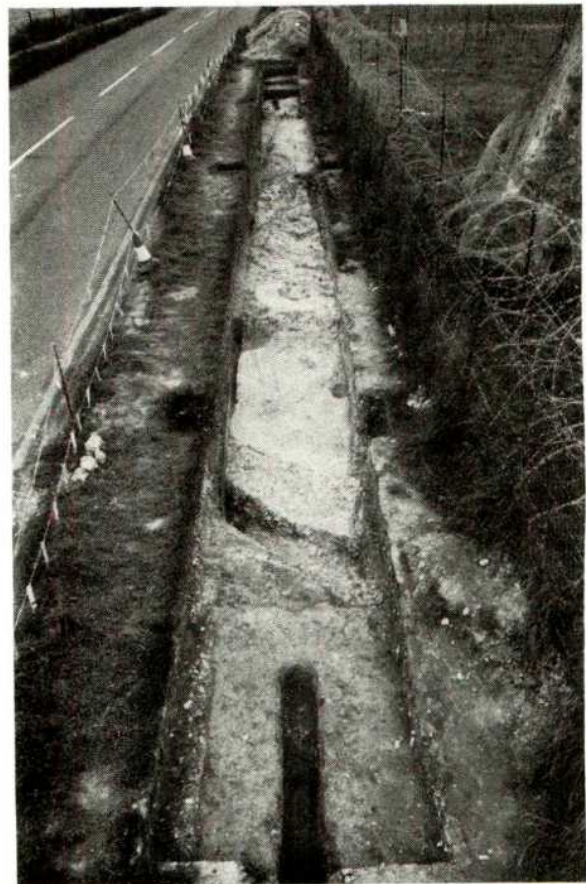


Fig. 2 Excavation in progress, looking south-east. The stone pit shows as a large depression in the chalk, truncated by the Heelstone ditch which crosses the trench in the foreground. Emerging from unexcavated ground to the right, is the smooth, crushed chalk surface created by the weight of the standing stone. The rectangular excavation at the bottom of the photograph shows the point reached by the Post Office cable when archaeological investigation began. Behind the coils of barbed wire (erected around the site for the period of the summer solstice festivities) can be seen the base of the Heelstone.

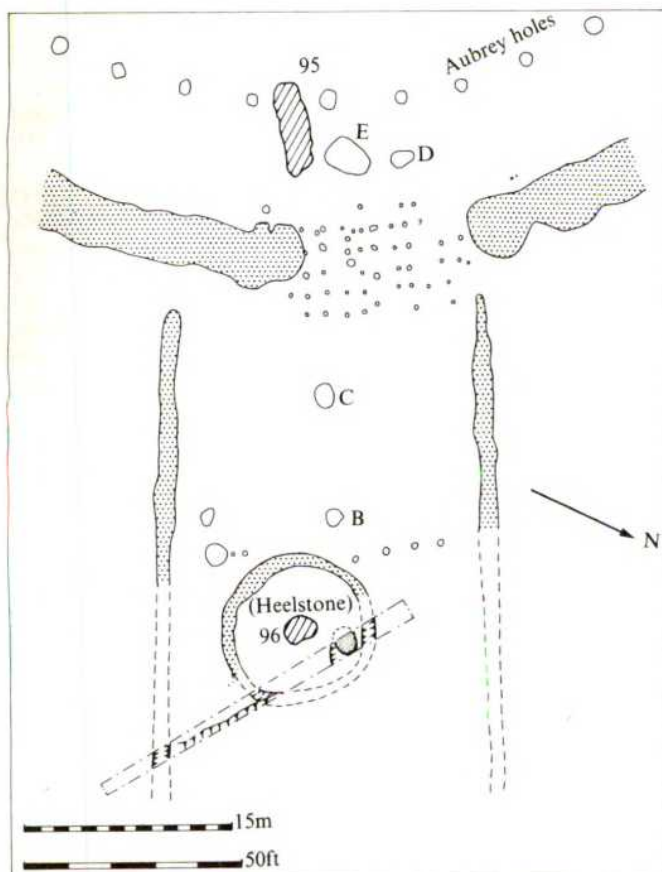


Fig. 1 Area around the entrance to the earthwork enclosure at Stonehenge, showing pits (letters), stones (numbers) and the trench excavated in 1979. The newly discovered pit is north of the Heelstone, and the stone impression is shown shaded.

Heelstone was excavated; this ditch had been refilled at the time the ditches and banks defining the Avenue were created, around 2170 BC (BM-1164: 1728 ± 68 b.c., HAR-2013: 1770 ± 70 b.c.; calibrated after Clark¹¹, 2140 ± 150 BC and 2200 ± 155 BC respectively).

The location of this previously unknown stone (called stone 97 in the sequence, after number 96, the Heelstone) questions the relationship of the Heelstone to the site's alignment. The eccentric position of the latter stone to the main axis of the site, which is approximately oriented on the midsummer sunrise has proved puzzling. If stone 97 was standing at the same time as the Heelstone, and the two were erected as part of a single plan, the problem may be solved, for the two stones sit astride the axis. Indeed, Newall predicted the presence of a stone beside the Heelstone on this hypothesis¹². The midsummer Sun would then rise on an alignment, accidentally or deliberately contrived, between the two stones. However, the Heelstone may have been erected after the now missing stone had been removed, with or without a period during which no stone was there at all. This interpretation was favoured by Atkinson (personal communication) who calculated the azimuth of the apparent centre of the impression left by stone 97 from the estimated centre of Stonehenge I as about $48^\circ 21'$. The precise location of the stone could not be determined from the excavated evidence—the stone's shape and size above ground are only definable within very broad limits.

Adopting the hypothesis of the Heelstone as one of a pair of stones, how does this pair fit into the overall arrangement of known stone locations at the site? The distance¹³ both between the main sarsen circle (specifically stones 1 and 30) and the pair of stones at the earthwork entrance (stone 95 and that represented by hole E), and the distance between this last pair and the Heelstone is some 30 m. Can we extend this pattern and interpolate an "imaginarie Walke of stones" similar to those at Avebury¹⁴?

The confused picture presented by antiquarian literature^{15,16} has prompted magnetic and resistivity surveys along part of the Avenue. The results show a strong disturbance in the chalk precisely where the next pair of stones should be, moving on from the Heelstone. The geophysical anomaly indicates the possible presence of only one pit. However, the model would predict this, because for both the known pairs the northwestern stone was removed in prehistory and the pit backfilled with chalk: it cannot be assumed that either of these pits would be visible to sub-surface detecting devices.

The interpretation of the newly discovered stone 97 as a partner for the Heelstone does not, however, exhaust all possibilities; Atkinson has other opinions. An alternative would be that the stone is part of a single row, of which two other stones were identified by Hawley. The now vacant pits (labelled by Atkinson as B and C) and the new pit are equidistantly spaced on a straight alignment. The apparently smaller size of holes B and C, however, suggests that the stones they contained (and note that the existence of these stones is largely hypothetical) did not match the dimensions of 97. Indeed, the new pit is so large, apparently having been sectioned by Atkinson *et al.* in 1956 (ref. 17), that it possibly consists of intersecting pits that fulfill both these suggested roles.

Four outlying features, in addition to the Heelstone, are frequently used in complex geometrical or astronomical constructions. Of these four 'Station Stones', only two survive in position; the others are evidenced by poorly described hollows in the chalk. Their importance lies in their location on the four corners of a perfect rectangle that is aligned on the site's major axis. However, evidence suggests both that the Heelstone and the Station Stones may have had companions, and that the basis of the latter's location was not so much a rectangle as a circle. For this interpretation, we must consider a ring that lies on the periphery of the site.

The 56 more or less regularly spaced pits in this ring were first known as the 'X holes' and later as the 'Aubrey holes'¹⁸, although the association of these 56 pits with the five 'cavities' observed by Aubrey is unconvincing, not least because of the discrepancy in numbers. Note also that these five hollows, supposedly representing the collapsed fill of pits levelled well over four millennia before Aubrey's visit, were apparently

invisible when Gale and Stukely were at Stonehenge less than 60 years later, to Cunnington and Hoare at the end of the eighteenth century, and certainly not to be seen by Hawley or Newall in this century. It therefore seems likely that if the cavities really were prehistoric in origin, whatever they contained had been removed not very long before Aubrey's visit. This being so, they could only have held standing stones.

Atkinson considered the possibility that the Station Stones were the remnants of a former circle, but has preferred to associate them with the large sarsen structure as four outliers, reserving "the bare possibility that they once formed part of an earlier setting of sarsens within the bank"¹⁹. The main evidence for this putative circle consists of three hollows on the site's perimeter—pits F, G and H. As Atkinson and his colleagues noticed, these pits are correctly located to form, together with the Station Stones and 13 hypothetical pits (mostly in unexcavated parts of the site), a circle of perhaps 20 more or less evenly spaced pits, two of which still retain stones, just outside the ring of X holes. A search was made for one of these pits in the 1950s (in the region of X holes 4 and 5) with a variety of sub-surface detecting devices, but nothing was found²⁰. However, it is possible that a pit does exist, but with a chalk fill rendering it effectively invisible without excavation (as would probably have been the case with the newly discovered pit). Hawley considered at least two of these hollows to be natural (one became known as the 'bush hole'), but this appearance could be given by pits with chalky fill exposed by an excavator with very limited knowledge of pedology and the complex geomorphology of the site. Finally, there was circumstantial evidence²¹ that a stone existed on the perimeter of this hypothetical circle near X hole 28.

The cavities which Aubrey observed exhibit both a coincidence in number with the above features, and a remarkable similarity in their respective locations (Fig. 3). The cavities match holes F to H, and Aubrey's two stones the Station Stones 91 and 93. The remaining cavities would correspond to holes now conventionally described as pits for Station Stones 92 and 94. The only major discrepancy lies in Aubrey's relative transposition of stone 93 and the hollow for 'stone' 94. Without further excavation, the existence of this stone circle must clearly remain an open question.

This consideration of the Station Stones and the Aubrey holes, like the description of the newly discovered stone 97, exemplifies an unavoidable, if inconvenient, problem concerning Stonehenge—we really know very little indeed of the nature and sequence of activities and structures at the site.

I thank staff and students of the Department of Archaeology, University of Southampton, and British Telecommunications for cooperation, and Professor Atkinson for interest and unpublished data. The geophysical surveys of the Avenue were conducted by the Department of the Environment's Ancient Monuments Laboratory, who also took plaster and silicone rubber casts of the stone pit bottom. The excavation was funded by the Department of the Environment through the Wessex Archaeological Committee Implementation Fund.

Received 9 October 1980; accepted 14 January 1981.

1. Hawkins, G. S. *Nature* **200**, 306–8 (1963).
2. Hawkins, G. S. *Nature* **202**, 1258–61 (1964).
3. Hoyle, F. *Nature* **211**, 454–6 (1966).
4. Newham, C. A. *Nature* **211**, 456–8 (1966).
5. Colton, R. & Martin, R. L. *Nature* **213**, 476–8 (1967).
6. Colton, R. & Martin, R. L. *Nature* **221**, 1011–12 (1969).
7. Robinson, J. H. *Nature* **225**, 1236–7 (1970).
8. Newton, R. R. & Jenkins, R. R. *Nature* **239**, 511–2 (1972).
9. Atkinson, R. J. C. *Nature* **275**, 50–2 (1978).
10. Atkinson, R. J. C. *Stonehenge* (Penguin, Harmondsworth, 1979).
11. Clark, R. M. *Antiquity* **49**, 251–66 (1975).
12. Newall, R. S. *Antiquity* **3**, 75–88 (1929).
13. Hoare, R. C. *The Ancient History of South Wiltshire*, 144 (William Miller, London, 1812).
14. Long, W. *Wiltshire Archaeol. nat. Hist. Mag.* **16**, 33 (1876).
15. Lukis, W. C. *Proc. Soc. Antiquaries* **7**, 2nd Ser., 268–71 (1877).
16. Newall, R. S. *Wiltshire Archaeol. nat. Hist. Mag.* **61**, 93 (1966).
17. Atkinson, R. J. C. *Stonehenge*, 203 (Penguin, Harmondsworth, 1979).
18. Hawley, W. *Antiquaries J.* **1**, 30 (1921).
19. Atkinson, R. J. C. *Stonehenge*, 78–79 (Penguin, Harmondsworth, 1979).
20. Atkinson, R. J. C. *Stonehenge*, 70 (Penguin, Harmondsworth, 1979).
21. Atkinson, R. J. C. *Stonehenge*, 33 (Penguin, Harmondsworth, 1979).

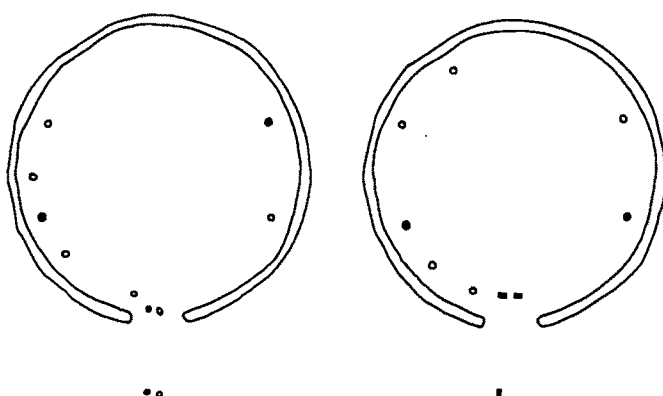


Fig. 3 An outer stone circle? Left: Holes F, G, H, 92 and 94 (○, working clockwise from the enclosure entrance) and the two Station Stones 91 and 93 (●); also shown are the locations of stone 95 and hole E (on the causeway) and stone 96 and the pit for 97 (outside the enclosure ditch). Right: the 'cavities' (○) and stones mapped by Aubrey. Enclosure ditch shown schematically.

Disappearance of common skate *Raia batis* from Irish Sea

K. Brander

Ministry of Agriculture, Fisheries and Food, Directorate of Fisheries Research, Fisheries Laboratory, Lowestoft, Suffolk NR33 0HT, UK

Records show that the common skate, *Raia batis*, has declined in abundance in the Irish Sea since the early years of the twentieth century, and is now very rare. As I report here, it is possible to calculate the highest mortality which the species will withstand without collapsing. It is likely that the mortality due to fishing has exceeded this level for some time and that the species will not recover while fishing continues. This represents the first clear case of a fish brought to the brink of extinction by commercial fishing.

In 1902 Herdman and Dawson¹ wrote that the common skate, *Raia batis*, was: "abundant in all parts (of the Irish Sea), and taken by line and by trawling all year round on nearly all our fishing grounds. The young are frequently taken in shrimp nets in shallow water". But the records of more than 800 trawl catches made by research vessels of the Ministry of Agriculture, Fisheries and Food (MAFF) throughout the Irish Sea during the past 10 yr show that no specimen of that fish was taken. None has appeared in trawl catches during research cruises by the Marine Science Laboratories, Menai Bridge during the past 15 yr (E. I. S. Rees, personal communication).

Enquiries among fishermen working in all parts of the Irish Sea indicate that *R. batis* was not uncommon in the late 1940s and that a few were taken during the 1950s and early 1960s, but there have been no recent records. In 1960, E. Hardy published a note in *Fishing News* on "The monster skate of Caernarvon Bay", giving details of a 9-foot-long specimen taken in September of that year. The Irish Specimen Fish Committee publishes an annual report, which in 1976 gave details of specimen-sized skate (more than 120 pounds) taken by anglers in Irish waters since 1956. Of those 20 specimen fish, 15 were from Strangford Lough. However, because of the decline in the skate populations, the Committee has suspended the species from its lists since 1976.

The identification of skate and ray species in the commercial landings in England and Wales is incomplete and they are not always recorded separately, however during the past 5 yr an average of less than 10 specimens of *R. batis* from the Irish Sea has been recorded annually at Fleetwood and these have all been from the North Channel. French commercial landings of *R. batis* from the area west of Scotland and Ireland are much higher, but here too the decline has been very rapid. Trawlers from Concarneau landed 922 tonnes of *R. batis* from this area in 1969, but only 85 tonnes in 1979. During the same period the landings of all other species of rays at Concarneau fell by only 15% (M. H. Du Buit, personal communication).

Several aspects of the skate's biology make it particularly difficult to protect from overfishing. Du Buit², commenting on the declining catches by French trawlers in the Celtic Sea, ascribed its vulnerability to slow growth, high age of maturity and low fecundity. Holden³ has modelled the population dynamics of other ray species from the Irish Sea, but did not distinguish between the mortality on immature and mature fish, and thus did not identify the crucial importance of fishing mortality on immature fish in determining the level of exploitation at which the stock would collapse. I have adapted Holden's method and have used his rates of egg laying to calculate the levels of mortality on immature and mature fish beyond which the stock will collapse.

For a population to remain in equilibrium, the mortality rate of mature fish (Z_m) must equal the net rate of recruitment of mature fish. The net recruitment is given by the number of eggs laid per female per year ($E/2$, because only eggs developing into

females are included) multiplied by the survival from egg laying to maturity ($e^{-Z_i t_m}$, if the immature mortality Z_i occurs evenly over the number of years t_m needed to reach maturity).

$$Z_m = (E/2) x e^{-Z_i t_m}.$$

The number of eggs laid per female depends on the size of the female and hence will vary because the average size of mature fish is affected by changing Z_m . For ray species, Holden⁴ has shown that the rate of egg laying is proportional to body length, not weight, so that there will be little change in $E/2$ as Z_m varies. The rate of egg laying by *R. batis* is not known but can be estimated as about 40 per year, using a relationship between fecundity and size at birth⁵. According to Du Buit², the age of first maturity for *R. batis* is 11 yr. The equation thus gives the relationship between Z_m and Z_i shown in Fig. 1. Any combination of values of Z_m and Z_i to the right of curve *a* will result in the collapse of the population. If the population was subject only to instantaneous natural mortality of say 0.2 before maturity, then a very high rate of exploitation would be possible on the mature stock, although this of course assumes that the direct relationship between number of mature females and net recruitment persists at all stock levels, which is unlikely. However, the fishing mortality on immature fish is probably the same as that on mature fish, because they are vulnerable to capture from birth by the same gear.

R. batis hatch from the egg at a length of ~22 cm, with a disk width of 13–15 cm and are retained by the mesh sizes in use. The line expressing the ratio $Z_m = Z_i$ is shown in Fig. 1 and it can be seen that the stock will collapse at a mortality above ~0.37. Even if the immature mortality were only half of the mature rate (that is $Z_m = 2 \times Z_i$, also shown in Fig. 1) the stock would still collapse when Z_i was 0.31. In fact, the immature mortality may well be higher than the mature mortality, because in addition to being caught in the same gear as the mature fish, the young skate were also taken frequently in shrimp nets in shallow water. Also, no account has been taken of egg and hatching mortality, which may be high and will act in a manner exactly analogous to reduced fecundity. Curve *b* in Fig. 1 expresses either 50% egg mortality or fecundity of only 10 female eggs per year instead of 20. This shows that reduced (or increased) fecundity has a relatively small effect on the mortality at which the stock collapses. It is the net rate of recruitment which is important, and this is determined largely by the cumulative mortality during the immature stage.

No estimates of mortality rates are available for the ray species of the Irish Sea, but they are likely to be of the same order as those on other demersal species, which range from ~0.4

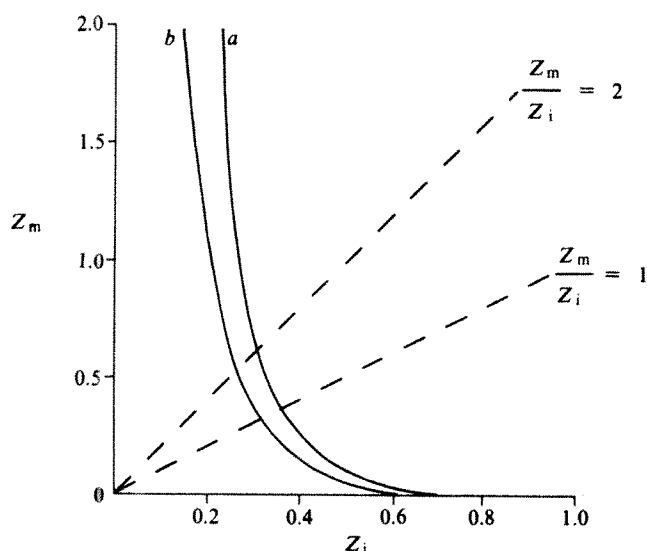


Fig. 1 Values of Z_m (mortality rate on mature fish) and Z_i (mortality rate on immature fish) to maintain the population in equilibrium. *a*, Fecundity = 20♀ eggs per yr; *b*, fecundity = 10♀ eggs per yr.

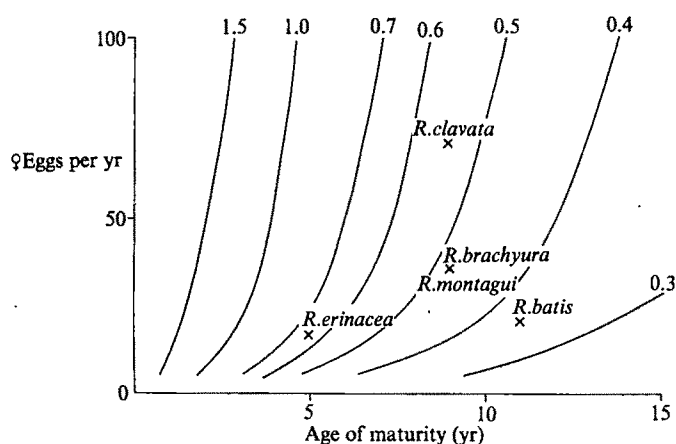


Fig. 2 Value of total mortality ($Z_i = Z_m$) beyond which the population will collapse, for different values of fecundity and age of maturity.

to 1.2 (total instantaneous rates on an annual basis⁶). The disappearance of *R. batis* is therefore not surprising, but that this took place virtually unnoticed or without comment in the fisheries literature, perhaps is surprising. Other ray species have also declined⁶, but are less vulnerable, as can be seen from Fig. 2. This shows the highest total mortality which the different species could withstand without collapsing, as a function of fecundity and age of maturity. It is assumed that $Z_m = Z_i$, but young *R. clavata*, *R. brachyura* and *R. montagui* may be less vulnerable to fishing mortality than *R. batis*, as they are smaller.

The prediction from Fig. 2 that *R. erinacea* would withstand the highest mortality of any of these species, in spite of having the lowest fecundity, emphasizes that net survival to maturity rather than fecundity *per se* governs the response to exploitation. For *R. batis* the only effective protection is probably a complete halt to all kinds of fishery in which it may be caught. Because this course of action is unrealistic, it has to be accepted that this species and others like it will be fished out as a consequence of the exploitation of other demersal fish.

Received 18 September 1980; accepted 13 January 1981.

1. Herdman, W. A. & Dawson, R. A. *Mem. Lancs. Sea Fish. Comm.* 11, 1-98 (1902).
2. Du Buit, M. H. J. *Cons. int. Explor. Mer* 37, 261-265 (1977).
3. Holden, M. J. in *Fish Population Dynamics* (ed. Gulland, J.) 187-215 (Wiley, London, 1977).
4. Holden, M. J. *J. Cons. int. Explor. Mer* 38, 110-118 (1975).
5. Holden, M. J. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer* 164, 360-367 (1973).
6. Brander, K. M. *Lab. Leaflet, MAFF Direct. Fish Res.*, Lowestoft, 36 (1977).

Infant-related influences on birth intervals in rhesus monkeys

M. J. A. Simpson, A. E. Simpson,
J. Hooley & M. Zunz

MRC Unit on the Development and Integration of Behaviour,
University Sub-Department of Animal Behaviour, Madingley,
Cambridge CB3 8AA, UK

That a mother's relationship with her infant could influence her subsequent reproductive history can be argued as follows. A rhesus monkey infant whose next sibling is to be born in a succeeding birth season¹ could compete with it by trying to postpone the date when its mother next conceives. If the mother preferred a shorter delay than the infant², processes of conflict³ and negotiation⁴ involving the two could show in some of the behavioural interactions constituting their relationship⁵. For example, a mother could try to hasten the onset of her next pregnancy by trying to promote more independent behaviour in her infant, perhaps by rejecting some of its bids for contact^{3,4}. We report here studies on our rhesus monkey colony and other colonies which support this argument.

We compared two reproductive groups of mothers in our colony¹: 'quick' mothers, who have an infant in the present birth season¹ and will conceive another in the coming mating season, with 'slow' mothers, who bear an infant this season but will not conceive another in the coming mating season. The two groups differed as follows: current infants of quick mothers included a higher proportion of males than did current infants of slow mothers (Fig. 1, $P < 0.01$ by Median test⁶, excluding two sons and four daughters whose mothers have not yet produced another full-term infant); current infants of quick mothers were more active, crossing grid lines in their cages more often⁷ during their first 12 weeks ($P = 0.048$, 0.008 and 0.035 in weeks 2, 4 and 8 respectively, Mann Whitney U-test); and quick mothers rejected^{7,8} their infants more often in seven of the eight observation weeks up to week 20 (Fig. 2, $P < 0.10$ in weeks 2-10, $P < 0.018$ in week 6 and < 0.011 in week 10).

In Bernstein's captive colony of rhesus monkeys⁹, more mothers of sons than of daughters were 'quick' (personal communication). Colvin (personal communication) found the same in the free-ranging rhesus population on Cayo Santiago in 1978-79 ($P = 0.0072$, Fisher exact probability test), but it has not been found in free-ranging populations studied over longer periods, either on Cayo Santiago (ref. 10, and Sade and Berman, personal communication) or on La Parguera (Leighton-Shapiro, personal communication). At this stage we presume that the special conditions in captive colonies of rhesus monkeys, and in

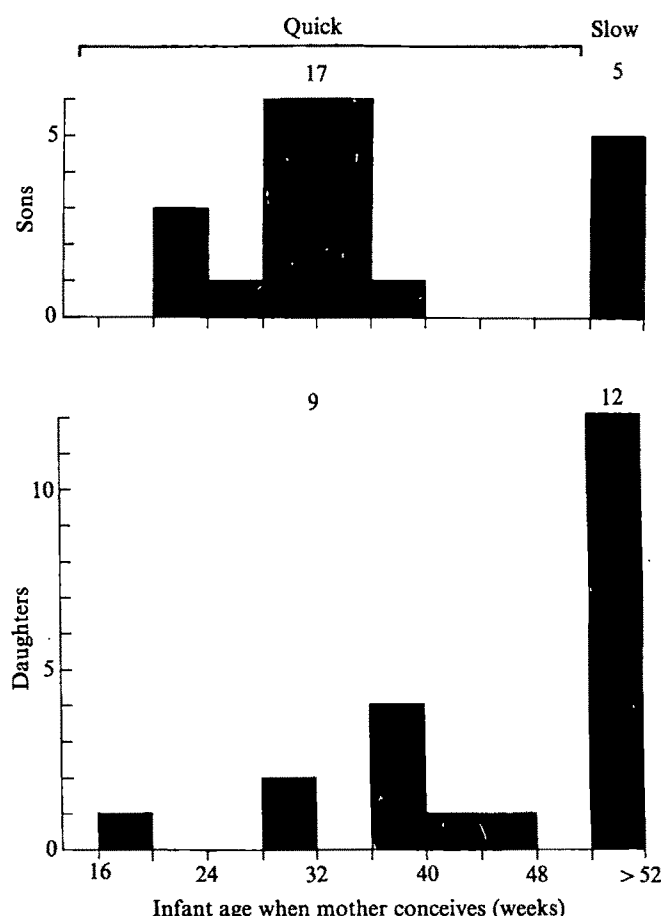


Fig. 1 Ages in weeks of current sons and daughters when their mothers conceive infants to be born in the next birth season (quick), and numbers of sons and daughters whose mothers fail to give birth in the next season (slow). The conception date of each infant to be born next season was calculated by subtracting 168 days²⁶ from its birthdate. The 43 infants and their mothers were healthy through at least the year following the births of the current infants, and those births occurring in the next season were full term¹. Our rhesus monkey colony has six separately caged groups, each with one adult male, 2-5 adult females, and immature individuals which have been born into the group¹.

some years in free-ranging populations, determine that an infant's sex becomes associated with its mother's next conception date.

In our colony, the tendency of mothers of males to conceive sooner could be explained in terms of infant activity level, if male infants as a group are more active. We found this to be the case in the infants' second week when more sons than daughters entered at least one new cage section during 6 h of observation ($P = 0.023$, Fisher exact probability test). Thereafter entry rates into new cage sections no longer differentiated male and female infants, but in all observation weeks during their first half year, males initiated more of the play episodes in which they were involved (see also ref. 11). (After week 2, when all infants can move freely, we do not expect males to enter more cage sections than females, because male social play often includes much rough-and-tumble activity in one place, whereas female play may include more chasing to and fro¹¹.)

If infant activity is associated with subsequent maternal conception date, this association should also show within the infant sex classes separately. Taking sons alone, those with quick mothers were more active in weeks 2, 4 and 6. Daughters with quick mothers were more active in all 10 observation weeks up to week 28 ($P < 0.022$ in weeks 8, 10, 16 and 20 by Mann Whitney U-test). Also, in all these weeks, daughters of quick mothers spent less time on their mothers' nipples ($P < 0.05$ in week 20, Mann Whitney U-test). If an infant's sucking behaviour is related to other aspects of its activity, one mechanism promoting early conception could involve sucking¹²⁻¹⁴ (at least in our mother-daughter dyads).

The association between maternal rejection scores^{7,8,15} and subsequent conception held true for sons and daughters separately throughout their first 10 weeks (Fig. 2), quick mothers being more rejecting. In a group of savannah baboons, less restrictive mothers had a shorter postpartum amenorrhea, although they did not conceive again sooner¹⁶. Baboon mothers who were less restrictive were also more rejecting^{4,17}. Associations between a mother's social position in her group and her fertility are also to be expected, and occur for our mothers of males (work in progress).

We provisionally suggest the following mechanism to explain why mothers who reject their infants should be more likely to conceive relatively early. A mother both responds to and tests the progress of her infant, conceiving earlier, the better its progress seems to be². That mothers could be affected by their infants' activity is suggested by the correlation between infant activity and maternal rejection rate in the infants' early weeks (Spearman rank-order correlation coefficient = 0.65, 0.80, 0.49 and 0.43 in weeks 2, 4, 6 and 8 respectively, $P < 0.007$). Mothers could even test their infants' capacity for activity by persisting in their rejecting behaviour until the infants begin to protest. The view of progress so acquired by the mother should be quite reliable, for by falsifying it² the infant could act to its own disadvantage. A clinging and sluggish infant would miss opportunities for play and exploration¹⁸.

This hypothesis does not explain why more mothers with male infants should subsequently be quick to conceive, unless we accept the greater activity of the sons at 2 weeks as a sufficient explanation. However, one or more of the following infant sex differences, expressed *in utero* and later, could also be effective. Thus our rhesus mothers pregnant with male fetuses received fewer threats, chases and attacks than those pregnant with females (work in progress), and pigtail macaque mothers (*Macaca nemestrina*) carrying male fetuses were less likely to require medical treatment (usually for bite wounds from social companions) than those carrying females¹⁹. Chamove's stump-tail macaques (*M. arctoides*), whose infants were removed at birth, bore subsequent infants sooner if they were male (personal communication). Rhesus mothers caged singly with their infants²⁰, and in our social groups (work in progress) played more often with male than with female infants, and male infants initiated a greater proportion of their play sessions than females¹¹.

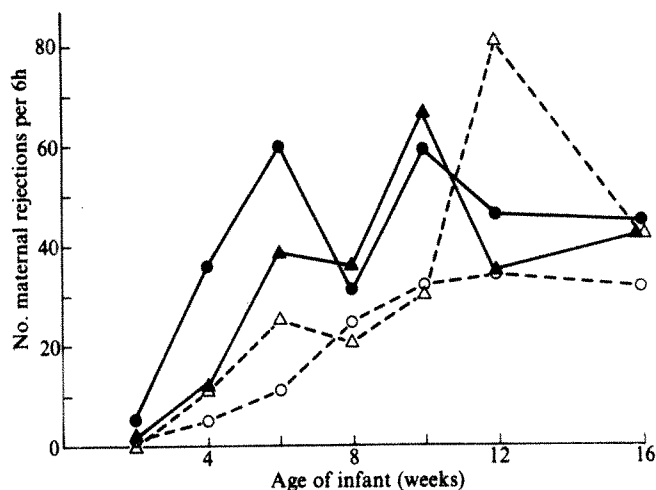


Fig. 2 Maternal rejection of daughters and sons through the infants' first 16 weeks, depending on whether mothers will conceive again between the infants' 16th and 52nd weeks (quick: ●, daughters; ▲, sons), or later (slow: ○, daughters; △, sons). Each data point is a median frequency of rejection from 6 h of observation¹⁵ of one of the groups. A rejection^{7,8,15} was scored each time a mother hit, pushed or bit her infant, blocked its access to her ventrum and/or nipple, or broke contact with it within 5 s of the infant initiating a period of physical contact with her. A rejection was also scored when the infant was already in contact with the mother's body and she blocked its access to her ventrum and/or nipple. The 43 infants were born to 20 mothers over a 7-yr period, and were treated as if they were statistically independent⁶. This is justifiable because we found that none of our six social groups, nor any of our families within social groups consistently produced infants that were male rather than female, or quick rather than slow.

The idea that mothers in good condition should produce sons rather than daughters²¹ suggests another explanation of our results. If old mothers were in poor condition and therefore 'slow', maternal condition could explain our observation that mothers of daughters tended to be slow. But many of our old mothers were vigorously maintaining high dominance status in their social groups, and were as quick or slow according to the sex of their infants as the remainder of our mothers.

High maternal rank could be the common factor underlying a tendency to have daughters⁴ and then to be slow. However, our high-ranking mothers of sons were among our 'quickest' mothers of all (work in progress).

The survival value to mother, and infant, of the mother being able to respond to and test her infant's progress is easily seen. It is less obvious, in terms of evolutionary theories, about how parents should invest care in their current and future offspring^{2,21-24}, why mothers of male rhesus infants, which are the slightly heavier sex at birth, should conceive again sooner. (Males are heavier in all studies known to us, but only statistically significant in that of Goy, personal communication.) In red deer (*Cervus elaphus*) the heavier stag calves are believed to be more costly to rear and in accordance with this idea their mothers are more likely to have a subsequent barren year²⁵. We wish to note the contrast provided by our rhesus monkeys and make a plea for comparable studies of more species.

The MRC supported the writing of this letter. The monkey colony was supported initially by the Mental Health Fund, and continuously by the MRC and the RS.

Received 18 August 1980; accepted 14 January 1981.

- Anderson, D. M. & Simpson, M. J. A. *Lab. Anim.* **13**, 275-281 (1979).
- Trivers, R. L. *Am. Zool.* **14**, 249-264 (1974).
- Hinde, R. A. *The Biological Bases of Human Social Behaviour* (McGraw Hill, New York, 1974).
- Altman, J. *Baboon Mothers and Infants* (Harvard University Press, 1980).
- Hinde, R. A. *Towards Understanding Relationships* (Academic, London, 1979).
- Siegel, S. *Nonparametric Statistics* (McGraw Hill, New York, 1956).
- Hinde, R. A. & Spencer-Booth, Y. *Anim. Behav.* **15**, 169-196 (1967).

8. Simpson, M. J. A. & Howe, S. *Behaviour* 72, 127-155 (1980).
9. Wilson, M. E., Gordon, T. P. & Bernstein, I. S. *J. med. Primat.* 7, 202-213 (1978).
10. Berman, C. M. *Anim. Behav.* 28, 860-873 (1980).
11. Harlow, H. F. in *Sex and Behaviour* (ed. Beach, F. A.) 234-265 (Wiley, New York, 1965).
12. Bonner, J., Franklin, M., Nott, P. N. & McNeilly, A. S. *Br. med. J.* 4, 82-84 (1975).
13. Delvoys, P., Badawi, M., Demagel, M. & Robyn, C. in *Progress in Prolactin Physiology and Pathology* (eds Robyn, C. & Harter, M.) 213-232 (Elsevier, Amsterdam, 1978).
14. Bowman, L. A., Dilley, S. R. & Keverne, E. B. *Nature* 265, 56-58 (1978).
15. Simpson, M. J. A. in *Social Interaction Analysis* (eds Lamb, M. E., Suomi, S. J. & Stephenson, G. R.) 137-156 (University of Wisconsin Press, Madison, 1979).
16. Altmann, J., Altmann, S. A. & Hausfater, G. *Science* 210, 1028-1029 (1978).
17. Altmann, J. in *The Development of Behaviour: Comparative and Evolutionary Aspects* (eds Burghardt, G. M. & Bekoff, M.) 253-277 (Garland STP11, New York, 1978).
18. Suomi, S. J. in *Minnesota Sym. Child Psychol.* 10, 201-228 (1976).
19. Sackett, G. P., Holm, R. A., Davis, A. E. & Farenbruch, C. E. in *Symp. 5th Congr. int. Primat. Soc.* (eds Kondo, S., Kawai, M., Ehara, E. & Kawamura, S.) 189-205 (Japan Science, Tokyo, 1975).
20. Mitchell, G. D. *Child Devel.* 39, 611-620 (1968).
21. Trivers, R. L. & Willard, D. E. *Science* 179, 90-92 (1973).
22. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B. G.) 136-179 (Aldine, Chicago, 1972).
23. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
24. Clutton-Brock, T. H., Guinness, F. E. & Albon, S. D. *Nature* 289, 487-489 (1981).
25. Guinness, F. E., Albon, S. E. & Clutton-Brock, T. H. *J. Reprod. Fert.* 54, 325-334 (1978).
26. van Wagenen, G. *J. med. Primat.* 1, 3-28 (1972).

Hepatitis B polypeptide vaccine preparation in micelle form

Jacinta Skelly, C. R. Howard & A. J. Zuckerman

WHO Collaborating Centre for Reference and Research on Viral Hepatitis, Department of Medical Microbiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

The immunoprophylaxis of hepatitis B is hampered by the lack of a technique for growing hepatitis B virus (HBV) in tissue culture. Plasma from persistently infected individuals¹, one source of viral antigen, contains characteristic 22-nm spherical particles which share a common antigen (the hepatitis B surface antigen or HBsAg) with the outer envelope of the 42-nm double-shelled DNA virus. Highly purified inactivated 22-nm particles have been shown to be safe and to confer protective immunity against HBV in a recent large-scale clinical trial². We have already described the extraction from the particles of a complex of two proteins which are antigenic determinants of HBV—the polypeptide with molecular weight (MW) between 22,000 and 24,000 (called p23) and the glycosylated polypeptide (called gp28) with MW in the range 26,000-29,000 which is thought to be the glycosylated form of p23. We now report the preparation from this complex of water-soluble protein micelles which may be a suitable basis for a second-generation hepatitis B vaccine.

In the preparation of viral vaccines, highly purified viral antigens may be significantly safer than other preparations. Many viral vaccines contain not only specific viral antigens but also components derived from the host tissues in which virus has been grown and against which those to whom the vaccine is administered may react. The availability of a chemically defined polypeptide viral antigen may be especially important for HBV, for there are several reports of the persistent presence in preparations of 22-nm spherical particles of host-serum proteins^{3,4} which may inhibit the immune response or even induce immunological side reactions⁵. Sources of HBV antigen such as the hepatoma cell line PLC/PRF/5 (ref. 6) and *Escherichia coli* containing cloned viral DNA (ref. 7) are similarly contaminated. The advantages of vaccines based on pure antigens have, however, been demonstrated with influenza, where a vaccine consisting of the neuraminidase and haemagglutinin subunits has been shown to be adequately immunogenic when given in two doses and less stimulating of side reactions than the whole virus vaccine.

In our search for a chemically characterized antigen of HBV, we have concentrated on the surface antigen (HBsAg) present in

the 22-nm particles in plasma, which consist of lipid and several viral and host proteins (including, for example, serum albumin^{8,9}). The two peptides p23 and gp28 (refs 8, 10), which form a 49,000-MW complex in non-reducing conditions¹¹, have identical amino acid sequences at the N-terminal end¹² and recent nucleotide sequences¹³ suggest that there is an internal sequence of 19 hydrophobic amino acids which may be the site of the intercalation of the proteins in the 22-nm particles. It seems reasonable, therefore, to consider gp28 and p23 as integral membrane proteins, by definition partially hydrophobic and insoluble in aqueous media in the absence of detergents¹⁴. Complete solubilization of the integral proteins of viral membranes, with minimal loss of serological activity, has been most successfully achieved by the use of non-ionic detergents such as Triton X-100^{15,16}. However, the polypeptides are poor immunogens in a monomeric form in detergent solutions¹⁷. Recently, a method of detergent removal has been described which allows the amphiphilic proteins to reassociate through their hydrophobic portions into a protein micelle structure, which remains water soluble due to the presence on its outer surface of the hydrophilic portions of the molecules¹⁸.

Micelles prepared from Semliki Forest virus membrane proteins were highly effective in inducing a protective response to the virus in mice¹⁷. Applying this type of strategy to 22-nm HBsAg spheres purified from the serum of a carrier chimpanzee, we have previously reported the extraction of a complex of gp28 and p23 by Triton X-100 solubilization and affinity chromatography⁸. The 3.9S complex of p23 and gp28 was prepared essentially according to this method⁸. Purified 22-nm spheres, labelled with ¹²⁵I, were disrupted with 2% Triton X-100. The complex was separated from other particle constituents by

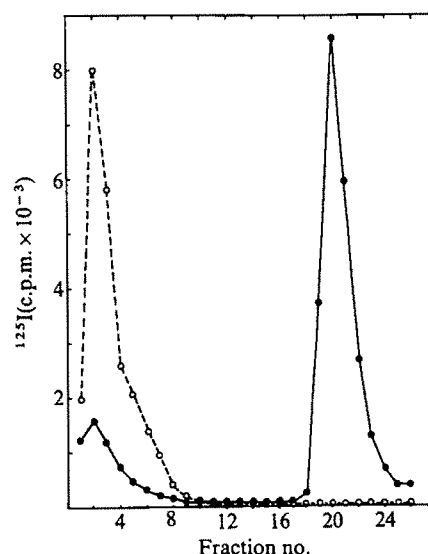


Fig. 1 Sedimentation of solubilized p23 and gp28 in 20-50% linear sucrose gradients in the absence (●) or presence (○) of 2% Triton X-100 throughout the gradient. The 22-nm HBsAg particles were purified from the serum of a persistently infected chimpanzee and radiolabelled with the Bolton and Hunter reagent, as previously described²². ¹²⁵I-HBsAg in 0.01 M Tris buffer was made 2% with Triton X-100 and 0.5 M with NaCl, and incubated overnight at 37 °C (ref. 8). The solubilized material was passed through a column of Con A-Sepharose equilibrated in the complete reaction buffer supplemented with 1 mM CaCl₂ and MnCl₂. The gp28-p23 complex was eluted with α -methyl mannoside and layered directly onto 20-50% linear gradients of sucrose in TNE buffer (0.50 M Tris-HCl, pH 7.4, 0.14 M NaCl, 0.001 M EDTA). After centrifugation at 220,000g for 24 h at 15 °C in a Beckman SW40 rotor, 0.5-ml fractions were collected from the top of the tube and counted to locate the position of the polypeptides. The direction of sedimentation is from left to right.

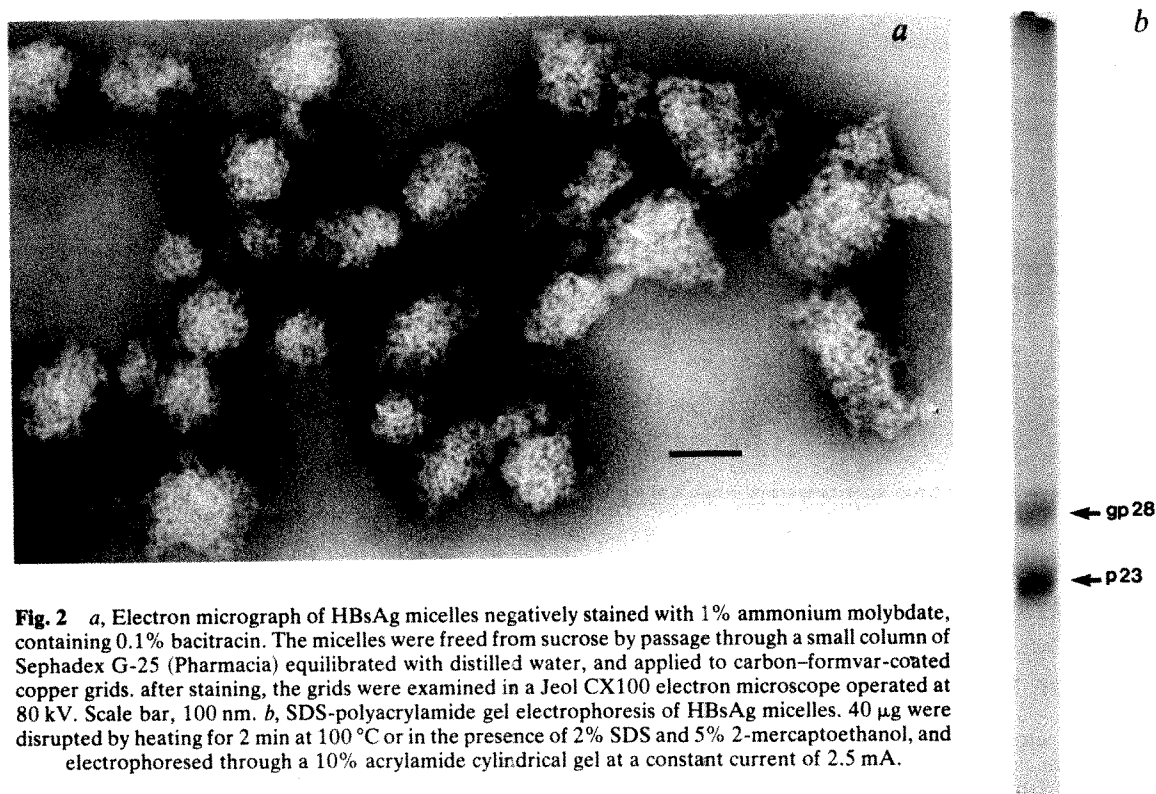


Fig. 2 *a*, Electron micrograph of HBsAg micelles negatively stained with 1% ammonium molybdate, containing 0.1% bacitracin. The micelles were freed from sucrose by passage through a small column of Sephadex G-25 (Pharmacia) equilibrated with distilled water, and applied to carbon-formvar-coated copper grids. After staining, the grids were examined in a Jeol CX100 electron microscope operated at 80 kV. Scale bar, 100 nm. *b*, SDS-polyacrylamide gel electrophoresis of HBsAg micelles. 40 μ g were disrupted by heating for 2 min at 100 °C or in the presence of 2% SDS and 5% 2-mercaptoethanol, and electrophoresed through a 10% acrylamide cylindrical gel at a constant current of 2.5 mA.

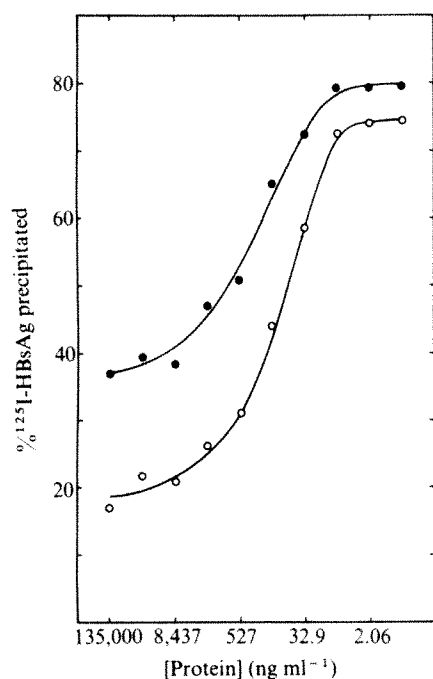


Fig. 3 Competitive radioimmunoprecipitation assay of HBsAg micelles (●) and intact 22-nm particles (○). Both samples were adjusted to 360 μ g ml⁻¹ of protein and serially diluted fourfold in phosphate-buffered saline (PBS) containing 0.5% of bovine serum albumin (BSA). A 150- μ l aliquot of each dilution was mixed with an equal volume of a 1:100 dilution of rabbit hepatitis B surface antibody (anti-HBs) (Hoechst) and incubated at 37 °C for 3 h. ¹²⁵I-HBsAg (intact 22-nm particles, 10 μ l, 24,000 c.p.m.) was then added and incubation was continued overnight at room temperature. A 150- μ l aliquot of protein A-Sepharose (Pharmacia) at a concentration of 5% in PBS-BSA was then added and the samples were shaken vigorously for 1 h at room temperature. The Sepharose was pelleted by low-speed centrifugation and the supernatants and pellets separated and counted. The results are expressed as the percentage of total counts associated with the precipitated immune complexes.

affinity chromatography on concanavalin A (Con A)-Sepharose in the presence of the detergent, and retained specific reactivity with anti-HBs after elution with α -methyl mannoside. The complex in detergent solution was then centrifuged into two 20–50% w/v sucrose gradients, one of which contained 2% Triton X-100 throughout (Fig. 1). In the gradient containing detergent, all the radioactivity remained at the top; in the absence of detergent a peak of ¹²⁵I-labelled protein was found two-thirds of the way down the gradient¹⁸. These results indicated that sedimentation of the complex into a detergent-free gradient separated the polypeptides from the detergent and allowed them to reassociate during centrifugation into a particulate protein micelle form with an increased sedimentation coefficient.

The buoyant density of the micelles in caesium chloride was 1.25 g ml⁻¹ compared with a density of 1.19 g ml⁻¹ for intact 22-nm particles, an increase consistent with the removal of most of the lipid by the solubilization procedure. Electron microscopy showed that the micelles were pleomorphic and fluffy in appearance, with diameters in the range of 60–200 nm (mean 120 nm) (Fig. 2a). Immune electron microscopy showed that the micelles could be both self-aggregated and co-aggregated with intact 22-nm particles by anti-HBs¹⁹. Polyacrylamide gel electrophoresis (Fig. 2b) showed that both gp28 and p23 were present in the same proportions as in the original detergent extract, indicating that the micelles were not formed preferentially from one or other polypeptide. Although the possibility that each polypeptide might have segregated into different micelles was not examined, it seems unlikely in view of the recent demonstration of a disulphide linkage between them¹¹. Gp28 and p23 together constituted 40% of the total protein of the purified 22-nm particles used as starting material in these experiments (see ref. 8). Ninety per cent of the gp28–p23 complex was recovered from the Con A-Sepharose column, and the final yield of the two polypeptides in micelle form was estimated to be 60–70% of the amount originally present in the intact 22-nm particles. The good recovery of total protein confirms the suitability of the method for large-scale production of a potential vaccine.

The degree of antigenic reactivity and specificity of the micelles was assessed by subjecting them to competition with

intact HBsAg particles for anti-HBs in a radioimmuno-precipitation test (Fig. 3). The micelles competed effectively with intact 22-nm forms, although the dilution curve had a slightly shallower slope than that obtained in a control experiment using unlabelled intact 22-nm particles as the competing antigen. Such a difference, indicative of altered protein conformation, was to be expected following removal of lipid and reassociation of the polypeptides into a purely protein form. It is difficult to give a precise estimate for the recovery of serological activity due to the difference in slope, but the competitive ability of the micelles was about 40% that of the intact 22-nm particle per weight of protein. The differences in surface area, and in surface distribution of antigenic sites between the two types of particles, may account for this discrepancy.

The immunogenicity of the micelles was compared with that of intact 22-nm particles in a mouse potency test, as it has been suggested that the serological response in mice may be a useful indicator of the immunogenic potential of candidate hepatitis B vaccines²⁰. Groups of six SWR/J mice were inoculated with three serial doses of 5, 10 or 20 μ g of each preparation in alum-adsorbed form, and the antibody levels in their sera measured by a protein A radioimmunoassay. The results obtained are shown in Fig. 4 in two forms. Figure 4a is a plot of the mean values of the serial dilutions of all sera obtained from mice inoculated with 10- μ g doses of micelles or intact particles.

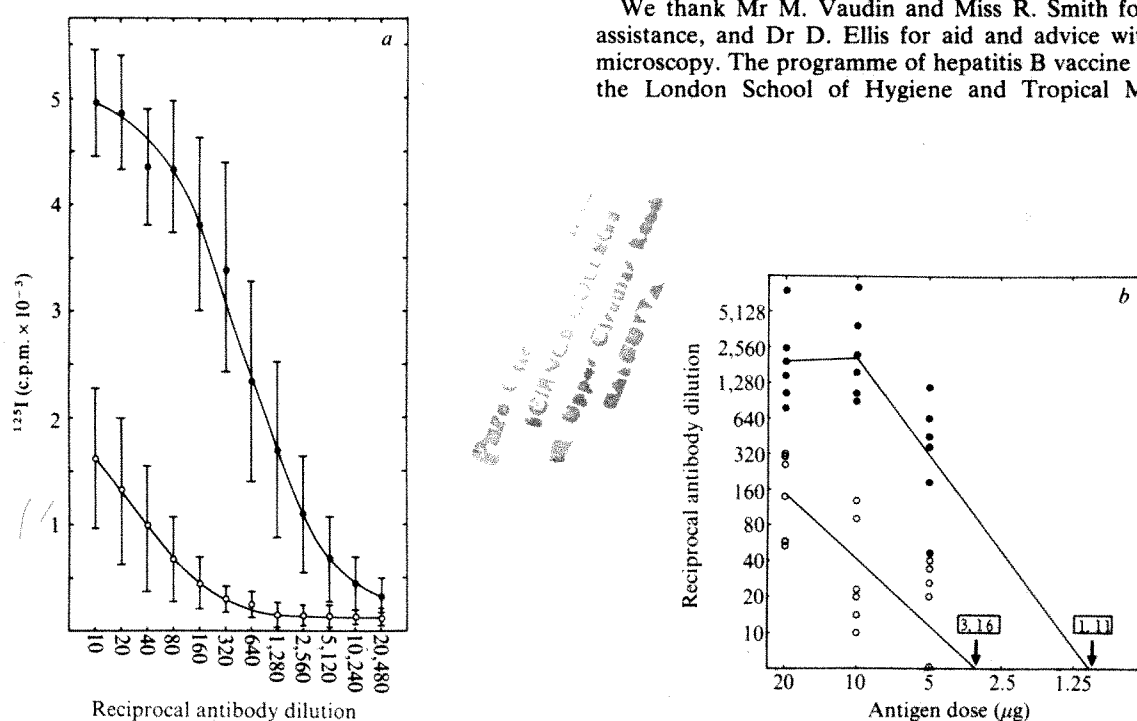


Fig. 4 Potency testing of HBsAg micelles and 22-nm particles. Both types of preparation were dialysed into 0.01 M phosphate buffer, pH 6.6, and adjusted to a final concentration of 200 μ g of protein per ml. A 1-ml aliquot of a 2% suspension of aluminium phosphate was added to 6 ml of each sample and the protein adsorbed to the gel by mixing the suspensions for 3 h. The alum was then centrifuged and washed once with PBS. Over 95% of the antigen was adsorbed by the alum. The alum was resuspended in PBS to a final concentration of 0.4%. The suspensions were then diluted appropriately with 0.4% alum in PBS so that both the volume and alum content of the inocula were the same for all antigen doses. Six mice per dose were injected intraperitoneally with 20, 10 or 5 μ g of either the intact 22-nm particle or the micelles in 100- μ l volumes. Two booster doses were given at weekly intervals. The mice were bled 11 days after the last inoculation and the titres of anti-HBs determined using a protein A radioimmunoassay. Flexible microtitre plates (Dynatech) were coated with 100 μ l per well of HBsAg at a concentration of 10 μ g ml⁻¹ in carbonate buffer, pH 9.5, by overnight incubation at 4 °C. The antigen was removed and 100 μ l per well of fetal calf serum added to block any remaining adsorption sites. After 2 h at 4 °C, the fetal calf serum was removed and the plates washed with PBS. 100- μ l aliquots of serial twofold dilutions of the test serum in PBS-BSA were incubated in the plate for 2 h at 37 °C, the wells were drained and rinsed and 100 μ l of ^{125}I -protein A was added to each. After 1 h at 37 °C, the plate was drained, washed several times with PBS and divided into individual wells for counting in a γ spectrometer. Protein A (Pharmacia) was iodinated by the chloramine-T method to a specific activity of 20–40 μ Ci μ g⁻¹. *a*, Protein A radioimmunoassay of anti-HBs in the sera of mice inoculated with 10- μ g doses of either micelles (●) or intact 22-nm particles (○). The sera were initially diluted 1:10 in PBS–0.5% BSA. The points represent the mean of six sera, and the vertical lines represent the standard deviations. *b*, Geometric mean antibody titres at different dose levels of HBsAg micelles (●) or 22-nm particles (○). Titres are expressed as the reciprocal of the dilution of antibody required to bind a given amount of radioactive protein A. Each point represents the titre in a single animal. The antigen extinction doses in each case are indicated.

Consistently higher levels of anti-HBs were induced in animals receiving micelles, possibly also with a higher affinity, as indicated by the steeper slope of the dilution curve compared with that of the sera from mice receiving intact particles, although direct affinity measurements were not made. This difference was seen at all dose levels tested. Figure 4b shows the antibody titres against dose for all animals, each point representing the titre of an individual serum. Extrapolation of the lines representing geometric mean titres of antibody gave antigen extinction values of 1.11 μ g and 3.16 μ g for HBsAg micelles and intact particles, respectively. Several factors may, singly or in combination, account for the greater immunogenicity of the micelles, including their large size, altered distribution of antigenic sites, and the absence in them of host-derived serum proteins such as albumin. The results presented here show that the antigenic complex of gp28 and p23 can be readily isolated in high yield from intact HBsAg 22-nm particles, and that this complex can be reassociated into a micellar form as has been described for other viral envelope proteins^{18,21}. The chemical purity, specific serological activity and powerful immunogenicity of the micelles, taken together with the ease of their preparation on a large scale, strongly favour their development as an alternative second-generation hepatitis B vaccine. This physical form may indeed also be suitable for vaccines prepared using HBsAg polypeptides produced by the expression of cloned hepatitis B virus DNA in *E. coli* or other hosts⁷. Experiments to test the immunogenicity, safety and protective capacity of HBsAg micelles in primates are in progress.

We thank Mr M. Vaudin and Miss R. Smith for technical assistance, and Dr D. Ellis for aid and advice with electron microscopy. The programme of hepatitis B vaccine research at the London School of Hygiene and Tropical Medicine is

supported by the Wellcome Trust, and in part by the Department of Health and Social Security and Organon BV, Oss, Holland.

Received 19 August; accepted 24 December 1980.

- Almeida, J. D., Zuckerman, A. J., Taylor, P. E. & Waterson, A. P. *Microbios* **2**, 117–123 (1969).
- Szmuness, W. *et al. New Engl. J. Med.* **303**, 833–841 (1980).
- Neurath, A. R., Prince, A. M. & Lippin, A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2663–2667 (1974).
- Burrell, C. J. *J. gen. Virol.* **27**, 117–126 (1975).
- Zuckerman, A. J. *Nature* **255**, 104–105 (1975).
- Macnab, G. M., Alexander, J. J., Lecatsas, G., Bey, E. M. & Urbanowicz, J. M. *Br. J. Cancer* **34**, 509–515 (1976).
- Pasek, M. *et al. Nature* **282**, 757–759 (1979).
- Skelly, J., Howard, C. R. & Zuckerman, A. J. *J. gen. Virol.* **44**, 679–689 (1979).
- Sanchez, Y., Ionescu-Matin, I., Hollinger, F. B., Melnick, J. L. & Dreesman, G. R. *J. gen. Virol.* **48**, 273–284 (1980).
- Shih, J. W.-K., Tan, P. L. & Gerin, J. L. *J. Immun.* **120**, 520–525 (1978).
- Mishiro, *et al. J. Immun.* **124**, 1589–1593 (1980).
- Peterson, D. L., Roberts, I. M. & Vyas, G. N. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1530–1534 (1977).
- Valenzuela, P. *et al. Nature* **280**, 815–819 (1979).
- Singer, S. J. & Nicolson, G. L. *Science* **175**, 720–731 (1972).
- Helenius, A. & Simons, K. *Biochim. biophys. Acta* **415**, 29–79 (1975).
- Helenius, A. & Söderlund, H. *Biochim. biophys. Acta* **307**, 287–300 (1973).
- Morein, B. *et al. Nature* **276**, 715–718 (1978).
- Helenius, A. & von Bonsdorff, C.-H. *Biochim. biophys. Acta* **436**, 895–899 (1976).
- Almeida, J. *et al. J. Virol. Meth.* (in the press).
- Gerety, R. J., Tabor, E., Purcell, R. H. & Tyeryar, F. J. *J. infect. Dis.* **140**, 642–648 (1979).
- Simons, K., Helenius, A., Leonard, K., Sarvas, M. & Gething, M. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5305–5310 (1978).
- Skelly, J., Howard, C. R. & Zuckerman, A. J. *J. gen. Virol.* **41**, 477–457 (1978).

Ethyl β -carboline carboxylate lowers seizure threshold and antagonizes flurazepam-induced sedation in rats

P. J. Cowen, A. R. Green & D. J. Nutt

MRC Clinical Pharmacology Unit, Radcliffe Infirmary, Oxford OX2 6HE, UK

I. L. Martin

MRC Neurochemical Pharmacology Unit, MRC Centre, Hills Road, Cambridge CB2 2QD, UK

The mammalian central nervous system possesses specific high-affinity binding sites for the benzodiazepines and considerable evidence suggests that these binding sites are the pharmacological receptors through which these compounds act^{1–6}. Recently⁷, ethyl β -carboline-3-carboxylate (β -CCE) has been identified in both human urine and rat brain. β -CCE may be closely related to the endogenous ligand for the benzodiazepine receptor—it shows an affinity for the receptor of the same order as that of clonazepam, one of the most potent benzodiazepines, and is the first non-diazepinoid structure to be identified with an affinity in the nanomolar range. Furthermore, it is selective for the benzodiazepine receptor⁷. Clinically and in animal studies, benzodiazepines have anti-convulsant, hypnotic and anxiolytic actions⁸. We have therefore investigated whether β -CCE exhibits any of these properties in rats. We report here that, in contrast to the benzodiazepines, β -CCE lowers seizure threshold and reverses the sedative effect of flurazepam. If β -CCE has a close structural relationship to the endogenous ligand, benzodiazepines may be antagonistic at the receptor site.

β -CCE was synthesized by condensation of L-tryptophan with glyoxylic acid, esterification of the product with ethanol-HCl and oxidation with chloranil. The purity of the final product was >95%.

The effect of β -CCE on seizure threshold in rats was determined by measuring the dose of bicuculline, infused through a tail vein, required to elicit a seizure⁹. The β -CCE was dis-

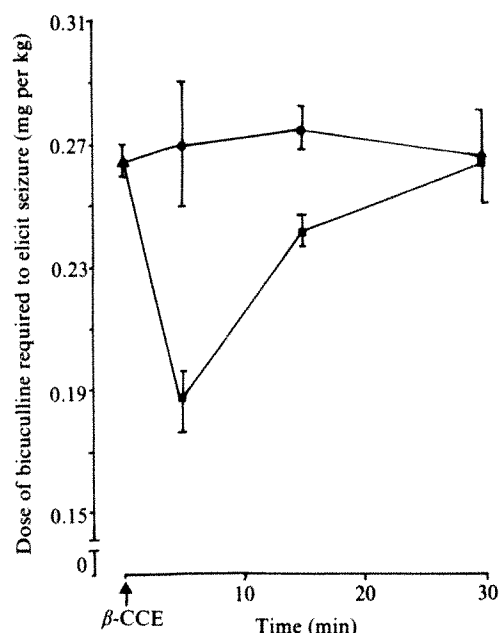


Fig. 1 Effect of β -CCE on the seizure threshold to bicuculline at various times following administration. Results show the seizure threshold for a group of un.injected rats (▲) and for groups of vehicle-injected (●) and β -CCE (5 mg per kg)-injected (■) animals. Results show mean $\pm$ s.e.m. for groups of four or more rats. The β -CCE group was significantly different from the vehicle-injected group— $P < 0.001$ at 5 min and $P < 0.005$ at 15 min.

solved in ethanol, diluted with propylene glycol (ethanol/propylene glycol, 1:2) and injected intravenously (i.v.). Control rats received the vehicle alone. Each rat was only used once for any determination.

When β -CCE was given i.v. at a dose of 5 mg per kg, there was a significant (31%) decrease in the dose of bicuculline required to elicit a seizure 5 min later. Only a small lowering of threshold was seen 15 min after administration and no effect was observed after 30 min (Fig. 1).

Pretreatment of rats with a dose of β -CCE as low as 0.5 mg per kg i.v. had a statistically significant effect in lowering seizure threshold 5 min later and the drug displayed a conventional log dose-response curve (Fig. 2). The effect of β -CCE on lowering seizure threshold plateaued at about 2.5 mg per kg, suggesting that the compound itself will not induce convulsions even at high doses.

When β -CCE (5 mg per kg) was given intraperitoneally, no effect on the bicuculline seizure threshold was seen 5 min later (vehicle 0.28 ± 0.02 mg per kg, $n = 4$; β -CCE 0.32 ± 0.04 mg per kg, $n = 4$). This could be due to rapid inactivation and may explain the short duration of action of the drug even when given i.v.

Benzodiazepine receptors are thought to be functionally linked to γ -aminobutyric acid (GABA) receptors^{10,11} and bicuculline is a GABA receptor antagonist¹². We therefore examined the effect of β -CCE on the seizure threshold using pentylenetetrazol (PTZ), as this convulsant agent also has GABA-antagonist properties¹³. We also investigated the effect on β -CCE on the seizure response to strychnine, a glycine antagonist, to see whether the action of β -CCE was specific to drugs acting on GABA functions. Five minutes after administration of β -CCE at a dose of either 5 or 1 mg per kg the seizure threshold to PTZ was lowered by a value (30%) comparable with that seen using bicuculline (vehicle 23 ± 3 (5); β -CCE (5 mg per kg) 16 ± 3 (6); β -CCE (1 mg per kg) 15 ± 3 (4), expressed as dose of PTZ (mg per kg) to elicit a seizure; β -CCE different from vehicle, $P < 0.001$, at both doses).

In contrast, no change was seen in the strychnine seizure threshold (vehicle 2.7 ± 0.2 (4); β -CCE (5 mg per kg) 2.7 ± 0.1 (4), expressed as dose of strychnine (mg per kg) required to elicit seizure). Following flurazepam (10 mg per kg) the PTZ seizure threshold is raised whereas the strychnine threshold is unaltered¹⁴.

When rats were given flurazepam (20 mg per kg intraperitoneally) they rapidly showed sedation and, in contrast to control animals, displayed very little 'exploratory' locomotor activity when placed in a new cage. This was reflected in low counts on Automex activity meters (Fig. 3). Injection of vehicle had no significant effect on this low level of activity. In contrast,

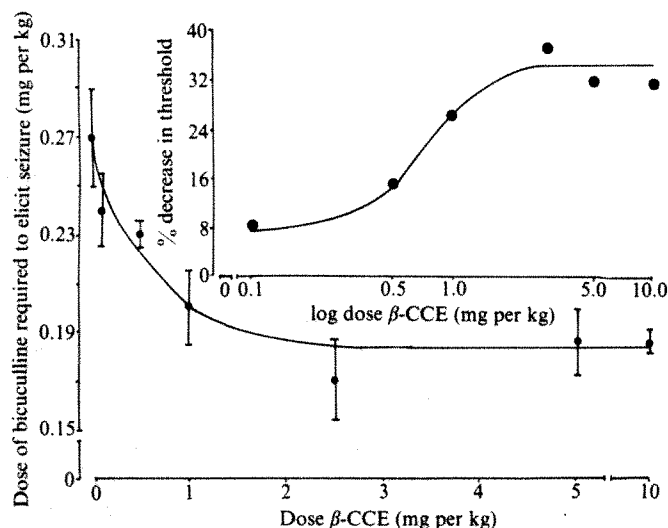


Fig. 2 Effect of dose of β -CCE on bicuculline-induced seizure threshold. Main graph shows dose of bicuculline required to elicit seizure 5 min after injection of various doses of β -CCE. All doses of drug were given in the same volume of vehicle (1 ml per kg). Zero value shows threshold following vehicle alone, which is not different from a saline-injected group (0.26 ± 0.01 , $n=4$). Different from vehicle control: 10, 5 and 2.5 mg per kg, $P < 0.001$; 1 mg per kg, $P < 0.01$; 0.5 mg per kg, $P < 0.02$; not significantly different: 0.1 mg per kg; four or more observations at each point. Inset shows the log dose-response curve of β -CCE measured against the mean percentage decrease in seizure threshold.

β -CCE (5 mg per kg i.v.) administration partially reversed the sedative effect of the flurazepam for about 10 min after injection (Fig. 3). β -CCE (5 mg per kg i.v.) alone did not produce an increase in locomotor activity compared with rats given vehicle. This suggests, therefore, that the increase in activity in flurazepam-treated rats receiving β -CCE resulted from an antagonism of the sedative effect of flurazepam.

There have been extensive investigations of the biochemical and behavioural effects of both natural and synthetic β -carbolines, some of which have been shown to have convulsant or anticonvulsant properties^{15,16}. However, of the several β -carbolines studied by Braestrup *et al.*⁷, only β -CCE and methyl β -carboline-3-carboxylate showed high-affinity binding (comparable with diazepam) to the benzodiazepine receptor. Although, as these authors admit, it is unlikely that β -CCE is the endogenous ligand, the ethyl ester group having been added during the isolation and purification process, they nevertheless suggest that it may be closely related to the endogenous ligand^{7,17}.

Sieghart and Karobath¹⁸ have recently demonstrated a heterogeneity of benzodiazepine receptors separable, after photoaffinity labelling, by SDS-polyacrylamide gel electrophoresis into two major bands with apparent molecular weights

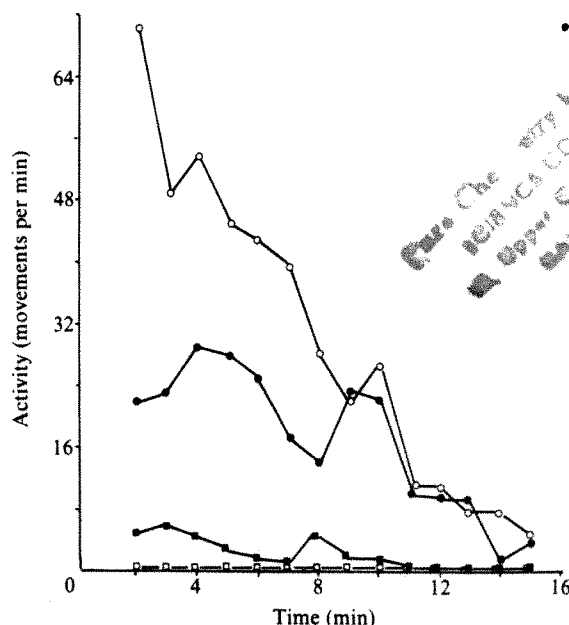


Fig. 3 The activity response of rats placed in cages on Automex activity meters following pretreatment with flurazepam (20 mg per kg, $n=4$, $\square$), flurazepam plus vehicle ($n=6$, $\blacksquare$), flurazepam plus β -CCE (5 mg per kg i.v., $n=7$, $\bullet$) or saline ($n=4$, $\circ$). Saline or flurazepam were given 15 min before the β -CCE or vehicle. Results show activity measured in recorded counts per min against time (min) after β -CCE or vehicle. Results at 2 min show response in the period 1–2 min post-injection. Flurazepam plus β -CCE is significantly different from flurazepam plus vehicle group at $P < 0.05$ or better for the points 2–10 min inclusive. Flurazepam plus vehicle was at no time significantly different from flurazepam alone.

of 51,000 (P_{51}) and 55,000 (P_{55}). The receptor subpopulations show a differential regional distribution, the cerebellum containing mainly P_{51} and the hippocampus approximately equal amounts of P_{51} and P_{55} . Data¹⁹ showing a lower IC_{50} for β -CCE in the cerebellum than hippocampus, indicates a greater affinity of β -CCE for the P_{51} than the P_{55} -subtype, but this difference in affinity is probably less than one order of magnitude. Such a small difference prevents a detailed interpretation of experiments, such as those reported here, performed on two different behavioural paradigms. We can conclude that the β -CCE has effects opposed to the benzodiazepines, but not that the reversal of the flurazepam-induced sedation and the convulsant effects of β -CCE are mediated through the same subpopulation of benzodiazepine receptors.

Received 23 October; accepted 31 December 1980.

1. Squires, R. F. & Braestrup, C. *Nature* **266**, 732–734 (1977).
2. Braestrup, C., Albrechtsen, R. & Squires, R. F. *Nature* **269**, 702–704 (1977).
3. Möhler, H. & Okada, T. *Life Sci.* **20**, 2101–2110 (1977).
4. Braestrup, C. & Squires, R. F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3805–3809 (1977).
5. Braestrup, C. & Squires, R. F. *Eur. J. Pharmacol.* **48**, 263–270 (1978).
6. Möhler, H. & Okada, T. *Science* **198**, 849–859 (1977).
7. Braestrup, C., Nielsen, M. & Olsen, C. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2288–2292 (1980).
8. Garattini, S., Mussini, E. & Randall, L. O. (eds) *The Benzodiazepines*, 685 (Raven, New York, 1973).
9. Nutt, D. J., Cowen, P. J. & Green, A. R. *Neuropharmacology* **19**, 1017–1023 (1980).
10. Tallman, J. F., Thomas, J. W. & Gallagher, D. W. *Nature* **274**, 383–385 (1978).
11. Guidotti, A., Toffano, G. & Costa, E. *Nature* **275**, 553–555 (1978).
12. Curtis, D. R., Duggan, A. W., Felix, D. & Johnston, G. A. R. *Nature* **226**, 1222–1224 (1970).
13. Simmonds, M. A. *Br. J. Pharmacol.* **70**, 75P (1980).
14. Nutt, D. J., Cowen, P. J. & Green, A. R. *Eur. J. Pharmacol.* (in the press).
15. Buckholtz, N. S. *Life Sci.* **27**, 893–903 (1980).
16. Ho, B. T. in *Current Developments in Psychopharmacology* Vol. 4 (eds Essman, W. B. & Valzelli, L.) 153–177 (Spectrum, New York, 1977).
17. Braestrup, C. & Nielsen, M. *Trends Pharmacol. Sci.* **1**, 424–427 (1980).
18. Sieghart, W. & Karobath, M. *Nature* **286**, 285–287 (1980).
19. Nielsen, M. & Braestrup, C. *Nature* **286**, 606–607 (1980).

Peripheral nerve grafts in hereditary leukodystrophic mutant mice (twitcher)

F. Scaravilli & Jean M. Jacobs

Department of Neuropathology, Institute of Neurology,
The National Hospital, Queen Square, London WC1N 3BG, UK

The twitcher mouse is a mutant affected by a form of leukodystrophy which shows close similarities to human globoid cell (Krabbe's) leukodystrophy. Transmission is by an autosomal recessive gene *twi*. Progressive loss of myelin sheaths from both central and peripheral nervous systems and the presence of inclusion-laden macrophages are characteristic findings. Morphological features of the twitcher have been described by Duchen *et al.*¹. Nerve iso- and allografting have been used to determine the roles of axon and Schwann cell in a number of mouse² and human³ nerve abnormalities. Schwann cells in a graft proliferate and become associated with regenerating host axons which grow through the graft into the host distal stump. In the twitcher, peripheral nerve axons do not degenerate but are thinner than normal, although there is considerable axonal degeneration in the central nervous system. In 15-day-old mutants, inclusions have been found in Schwann cells associated with apparently normal myelin sheaths. Grafting experiments might show whether the phenotype of this mutant is fully expressed in the Schwann cell, or if axons are also involved. In previous experiments, survival of transplanted Schwann cells was achieved by the use of T cell-suppressed^{2,3} or nude mice⁴. We report here that a twitcher nerve transplanted in immunologically unsuppressed animals reproduces all the characteristic features of leukodystrophy and conversely that Schwann cells from unaffected mice can produce normal myelin when associated with twitcher axons.

Myelin seems to form normally until the mice are ~15 days old although the myelin sheaths and axons are rather thinner than those of normal littermates. Later, the myelin sheaths start to degenerate and macrophages containing inclusions appear. By 3–4 weeks, peripheral nerves are macroscopically swollen; many axons are demyelinating or bare of myelin and others have thin myelin sheaths indicative of remyelination; the fibres are widely separated by endoneurial fluid and by large numbers of macrophages. There are various inclusions in macrophages and Schwann cells—those of particular interest are the multi-angular bodies and twisted tubules (Fig. 1a, b) also found in the human⁵ and canine⁶ forms of globoid leukodystrophy. Deficiency of the enzyme galactosylcerebrosidase was found in human cases of Krabbe's disease^{7,8} and also more recently in the twitcher mouse⁹, establishing the validity of the mutant as a model for the human disease.

Four twitcher litters (aged 40–45 days) were used, including normal and affected mice, and sciatic nerve grafts were carried out between littermates to limit as much as possible any immunological differences. Normal littermates which received twitcher transplants, originally used as controls, were expected to reject their grafts but in most cases they showed no signs of rejection. Grafts were also made from normal mice into twitchers but the life expectancy of the twitcher is so short that transplantation was carried out on 7-day-old littermates. At this age mutants were indistinguishable from their normal littermates and grafts were done at random. Successful grafts of normal sciatic nerves into twitcher sciatic nerves were achieved in two cases which survived for 35 and 37 days.

A total of 25 animals (20 normal and 5 twitchers) were grafted successfully. The results of twitcher sciatic nerve transplants into normal recipients 1–6 months after grafting are reported as well as those of the reverse experiment with 35- and 37-days

survival. To exclude the possibility of migration of Schwann cells from the proximal and distal stump into the transplant, two normal mice had the sciatic nerve severed at mid-thigh. After 48 h, at the peak of Schwann cell mitosis¹⁰, ³H-thymidine (4 µCi per g body weight) was injected intraperitoneally in each animal. After a further 24 h the two mice received a transplant from a twitcher nerve. Six months after the grafting they were perfused and blocks from proximal and distal nerve as well as from the transplant were examined using phase and electron microscopy.

One month after the operation the twitcher nerve transplant seemed to be well vascularized; many axons were myelinated and there was no sign of demyelination; Schwann cells contained no inclusions but some had abnormally watery cytoplasm with scattered organelles and the endoneurial space was greater than normal. No macrophages or inflammatory cells were seen in or around the graft. In the distal stump, axonal regeneration and remyelination followed the normal expected pattern throughout the experiment. In the transplant, endoneurial oedema was more pronounced at 2 months (Fig. 2); some fibres showed early demyelination and numerous axons were surrounded by only a Schwann cell containing myelin debris; abnormally thin myelin

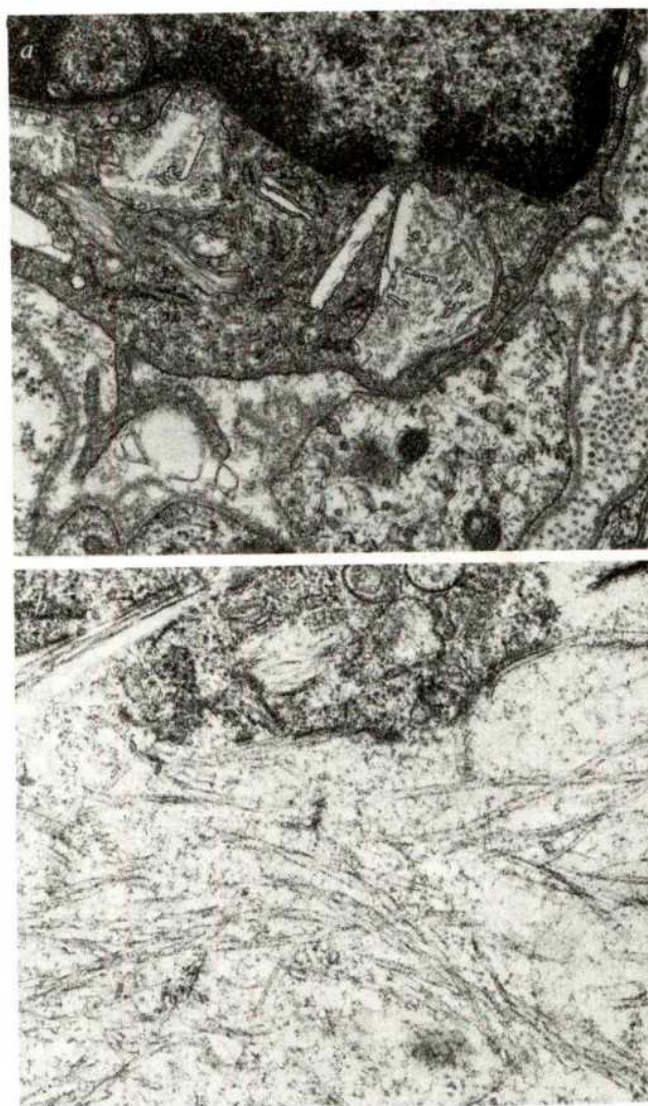


Fig. 1 Electron micrographs of macrophages containing a, multi-angular bodies and b, tubules, some of which appear twisted. Similar inclusions are also found in human and canine globoid cell leukodystrophy and are considered to be characteristic of the disease. a $\times 19,000$; b $\times 36,000$.

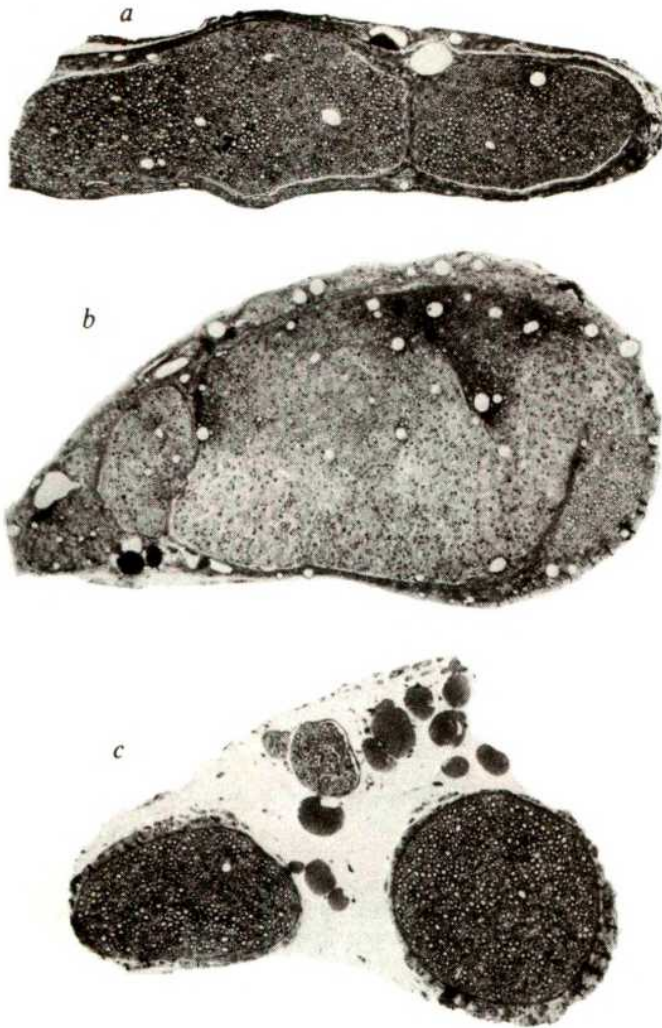


Fig. 2 One- μ m Araldite sections of: *a*, the normal proximal region; *b*, the twitcher graft; and *c*, the distal stump 2 months after transplantation. The graft is almost double the size of the proximal nerve due to the increased cellularity, severe endoneurial oedema and abundant collagen. *a*, *b*, *c* $\times 90$.

sheaths were indicative of remyelination (Fig. 3*a*, *b*). Twisted tubules, multi-angular bodies and other inclusions were seen in Schwann cells and macrophages.

At 3, 4 and 5 months some demyelinated axons were still present, but most were surrounded by apparently normal myelin and only a few Schwann cells contained myelin debris and specific inclusions. At 6 months many fibres appeared normal although the myelin sheath thickness was not comparable with the fibres in the normal distal stump. However, at each interval some features characteristic of the twitcher were still present; these comprised specific inclusions in Schwann cells and in macrophages, an occasional demyelinated axon and some endoneurial oedema (Fig. 4*a*, *b*). Autoradiographic studies confirmed the presence of labelled Schwann cell nuclei in the distal stump but no labelled cells were found in the transplant. The normal nerve grafted into the twitcher sciatic nerve showed the opposite trend, with larger myelinated fibres in the transplant than in the proximal or distal stump. These grafted nerve fibres appeared quite normal and there was no evidence of any of the other pathological features found in the twitcher host.

These experiments show that a twitcher nerve transplanted into a normal recipient reproduces all the characteristic features of leukodystrophy and, conversely, that Schwann cells from unaffected mice can produce normal myelin when associated with twitcher axons. It can be concluded, as in analogous studies by Aguayo *et al.*^{2,3}, that donor Schwann cells are able to

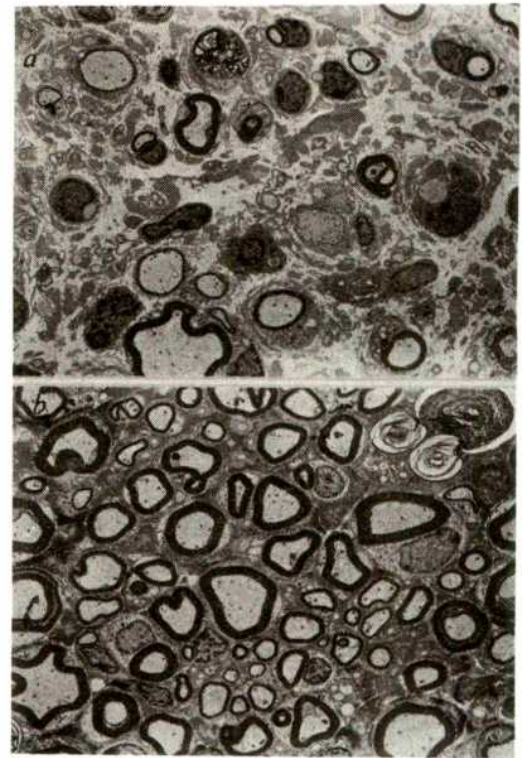


Fig. 3 Electron micrographs showing *a*, the transplanted twitcher nerve and *b*, the distal stump after 2 months. In the twitcher nerve, one fibre is demyelinating and several are completely demyelinated. In a few fibres myelin appears normal. Fibres are widely separated by endoneurial oedema, excess collagen and macrophages. $\times 1,500$.

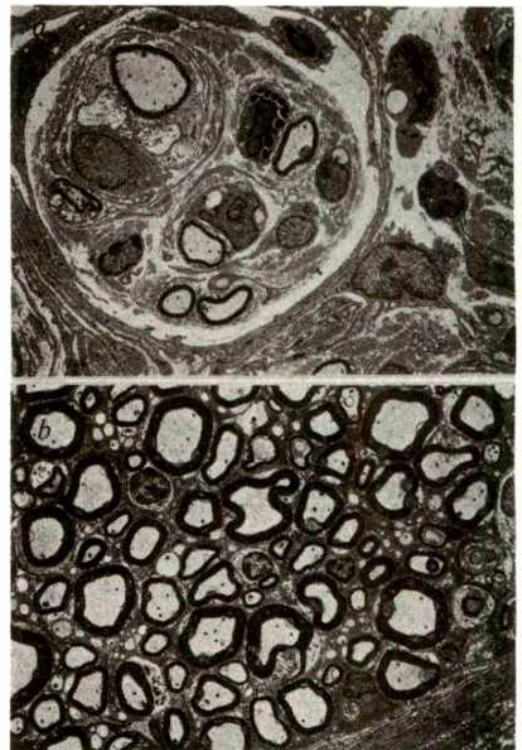


Fig. 4 Electron micrographs showing *a*, the transplanted twitcher nerve and *b*, the distal stump after 6 months. Several myelinated fibres are seen with sheaths thinner than those in the distal stump. Also visible are several inclusions in Schwann cells, a demyelinating and a thinly remyelinating fibre. The structures are more widely separated than those in the distal stump. $\times 1,500$.

reproduce an abnormality in a normal host, in association with normal axons. The twitcher axon has no apparent role in the development of the myelin sheath abnormality. Repopulation of the graft by Schwann cells from the host can be excluded on three grounds: (1) even after 6 months some specific abnormalities and abnormal myelin sheath parameters were found in the twitcher nerve grafted into a normal nerve; (2) there were neither specific inclusions nor abnormal Schwann cells and macrophages in the 37-day normal nerve transplanted into the twitcher; (3) there was no autoradiographic evidence of Schwann cell migration.

A puzzling feature was the apparent improvement towards a more normal situation in the 3–6-month grafts, which raises the question of the possible entry into the graft of enzyme from the normal host. A transitory improvement has been claimed by some authors in cases of Gaucher's disease (another enzyme deficiency condition) after transplantation of a normal spleen¹¹ and kidney¹².

The absence of rejection of the transplant without immunosuppressive treatment is unusual and does not agree with previous findings^{2–4}. It is known that a nerve transplant kept viable by anti-lymphocyte serum is quickly rejected when treatment is withdrawn³. The twitcher gene is maintained in a non-inbred strain¹ and tolerance of grafts would not be expected. The number of successful grafts in our experiment was too high to be accounted for by the random selection of immunologically matched littermates. Furthermore, a hypothetical immunological deficit in homozygous twitchers would not explain why littermates too, accept the grafts. It is possible that, although the close kinship plays a part, the non-inbred strain in which twitcher is maintained is of low immunological responsiveness¹³.

We thank Mrs J. Grainger for electron microscopic preparations. This research was supported by the Brain Research Trust and the National Fund for Research into Crippling Diseases.

Received 13 October; accepted 8 December 1980.

1. Duchen, L. W., Eicher, E. M., Jacobs, J. M., Scaravilli, F. & Teixeira, F. *Brain* **103**, 695–710 (1980).
2. Aguayo, A. J., Attwell, M., Trecarten, J., Perkins, S. & Bray, G. M. *Nature* **265**, 73–75 (1977).
3. Aguayo, A. J., Kasarjian, J., Skamene, E., Kongshavn, P. & Bray, G. M. *Nature* **268**, 753–755 (1977).
4. Dyck, P. K., Lais, A. C. & Low, P. A. *Neurology* **28**, 261–265 (1978).

5. Schochet, S. S., McCormick, W. F. & Powell, G. F. *Acta neuropath.* **36**, 153–160 (1976).
6. Yajima, K., Fletcher, T. F. & Suzuki, K. *J. neuropath.* **33**, 179–197 (1977).
7. Suzuki, K. & Suzuki, Y. *Proc. natn. Acad. Sci. U.S.A.* **66**, 302–309 (1970).
8. Malone, M. *Trans. Am. Soc. Neurochem.* **1**, 56 (1970).
9. Kobayashi, T., Scaravilli, F. & Suzuki, K. *Proc. 1st Symp. Neurological Mutations Affecting Myelination*, Seillac (Abstr.) 257–262 (Elsevier/North-Holland, Amsterdam, 1980).
10. Romine, J. S., Aguayo, A. J. & Bray, G. M. *Brain Res.* **98**, 601–606 (1975).
11. Grooth, C.-G. *et al.* *Lancet* **i**, 1260–1264 (1971).
12. Grooth, C.-G. *et al.* *Transplant Proc.* **11**, 1218–1219 (1979).
13. Guttman, R. D. *Transplant Proc.* **11**, 1622–1625 (1979).

Functional nuclei of LPS-nonresponder C3H/HeJ mice after transfer into LPS-responder C3H/HeN cells by cell fusion

Takeshi Watanabe & Junichi Ohara

Department of Immunology, Saga Medical School, Nabeshima, Saga, Japan

Splenic lymphocytes of C3H/HeJ mice are defective in responding to the mitogenic activity of bacterial lipopolysaccharide (LPS)^{1–4}, which has been attributed to a mutation at a single gene locus on chromosome 4 (ref. 5). It was suggested that C3H/HeJ mice lack a cell-surface receptor necessary for LPS recognition^{6,7}. However, specific binding of ¹²⁵I-labelled lipid-A of LPS to B cells was found to be equivalent in LPS-responder and LPS-nonresponder C3H/HeJ, C57BL/10 ScCR and C57BL/10 ScN mice⁸. This suggests that the genetic defect in C3H/HeJ mice is related to triggering mechanisms rather than to LPS binding. We recently established a method for separating mouse splenic lymphocytes into karyoplasts (minicells) and cytoplasts (enucleated cells) with high purity⁹. This method makes possible the exchange of nuclei from one type of lymphocyte to another. In the present study, karyoplasts were purified from LPS-nonresponder C3H/HeJ mice and transferred to mitomycin C-treated or X ray-irradiated splenic lymphocytes taken from LPS-responder C3H/HeN mice by fusion with polyethyleneglycol. These reconstituted hybrid cells could respond to LPS, suggesting that nuclei from C3H/HeJ mice could be activated by LPS if proper signals were provided from cell-surface receptors. Moreover, responsiveness to concanavalin A (Con A) of mitomycin C-treated or X ray-irradiated lymphocytes was also recovered by the microinjection of karyoplasts from untreated lymphocytes.

Cytochalasin B is known to induce enucleation of cell monolayers through centrifugation without inhibition of protein, RNA or DNA synthesis. Furthermore, the effects of this

drug on the cell mobility, cell form or permeability are readily reversible if it is removed. Enucleation can also be carried out with mouse L cells in suspension by centrifugation in Ficoll density gradients containing cytochalasin B¹⁰. We applied the enucleation method with Ficoll density gradient centrifugation to normal murine lymphocytes (see Table 1 legend; ref. 9). Enucleation occurred almost completely after centrifugation of splenic lymphocytes in Ficoll gradients containing cytochalasin B. Karyoplasts (minicells) were further purified by 1g velocity sedimentation in gradients of bovine serum albumin. After these procedures, staining with May-Giemsa's solution revealed that contamination of cytoplasts in this fraction was <1% and no intact cells were observed. Purity of karyoplasts was also examined by measuring protein and DNA contents as well as synthesis of protein, RNA and DNA after stimulation with mitogen (LPS, Con A).

In the present study, karyoplasts were prepared from LPS-nonresponder C3H/HeJ (HeJ) mice by the above method and fused to mitomycin C-treated spleen cells of responder strain C3H/HeN (HeN). Fusion was achieved with 50% polyethylene glycol containing poly-L-arginine. The addition of poly-L-arginine greatly enhanced the fusion efficiency and ~10–20% of splenic lymphocytes were found to contain two nuclei. Viability of cells was the same before and after fusion. Neither spleen cells nor karyoplasts obtained from HeJ mice responded to LPS in the range 1–100 µg ml⁻¹ (Table 1). Mitomycin C-treated HeN spleen cells fused with HeJ karyoplasts showed a strong DNA synthetic response to LPS whereas mitomycin C-treated cells or HeJ karyoplasts did not. Peak responses to LPS were seen on day 3 with fused and intact cells (data not shown). Mitomycin C-treated HeJ cells fused with HeJ karyoplasts showed no LPS mitogenic response, but they responded to Con A. In simple mixtures of mitomycin C-treated HeN spleen cells and HeJ karyoplasts, LPS responses were not observed.

In subsequent experiments, we used irradiated (1,200 rad) spleen cells as recipients instead of mitomycin C-treated cells—similar results were obtained (Fig. 1). Reconstitution of irradiated HeN spleen cells with HeJ nuclei by polyethylene glycol-mediated fusion resulted in a significant DNA synthetic response to LPS. This did not occur when HeJ karyoplasts were

inactivated by irradiation before transfer into irradiated HeN cells, suggesting that DNA synthesis in the fused cells was due to the injected HeJ nuclei rather than the host HeN cell nuclei. Fusion of nontreated HeJ cells with HeN karyoplasts did not restore LPS responsiveness although such cells showed strong responses to Con A. These data suggest that microinjection of nuclei into cells whose own nuclei are inactivated restores DNA synthetic responses to various mitogens and that the nuclei of nonresponder strain HeJ can respond normally to LPS if stimulatory signals from cell-surface receptors are appropriate.

The C3H/HeJ defect in LPS responsiveness has been attributed to a mutation of a single gene on chromosome 4 which maps close to the Mup-1 gene¹¹. Two contradictory explanations of this abnormality have been suggested—one that the mice lack suitable cell-surface receptor which is deficient in all LPS-nonresponder strains^{6,7} and the other that the LPS triggering signal(s) is not properly transferred from cell-surface receptors to the nuclei. The latter hypothesis is based on the findings that ¹²⁵I-labelled lipid-A of LPS bound equally well to the cell surfaces of responder and nonresponder mice⁸. Another possibility is that abnormal events occur in the nucleus in response to this particular mitogen in C3H/HeJ mice. In the present study we found that LPS nonresponsiveness is attributable to cell-surface or cytoplasmic events and that nuclei are functionally intact in this nonresponder strain.

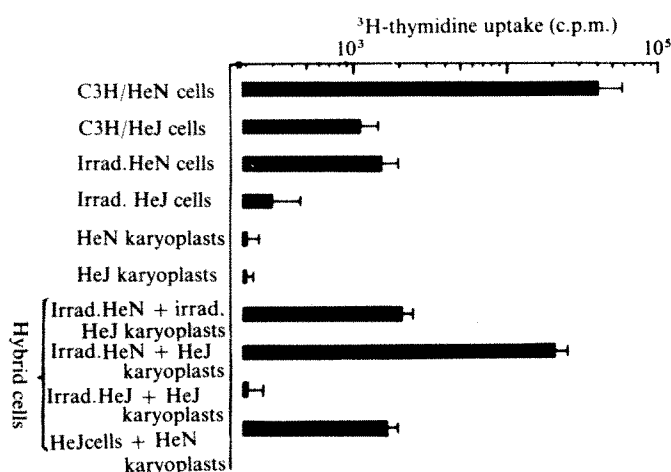


Fig. 1. Restoration of LPS responsiveness of irradiated C3H/HeN spleen cells by cell fusion. Preparations were made as described in Table 1 legend except that HeN or HeJ cells were inactivated by X-ray irradiation (1,200 rad) before fusion.

Since Carter reported¹² that the mould metabolite cytochalasin B caused nuclear extrusion of mouse L cells, this drug has been utilized to obtain large numbers of enucleated cells (cytoplasts) and minicells (karyoplasts). However, the methods have been generally limited to the cells, mainly cytoplasts, that adhere to glass or plastic coverslips after centrifugation. Wigler and Weinstein¹⁰ modified this enucleation method for use with mouse L cells in suspension. We demonstrated that separation of normal splenic lymphocytes into cytoplasts and karyoplasts without affecting cell viability, cell form and cell metabolism was possible using cytochalasin B in Ficoll gradients (Table 1). With our method, large numbers of pure karyoplasts are obtainable from normal splenic lymphocytes.

Fusion with polyethylene glycol¹³ has been extensively used for hybridization of lymphocytes and especially for monoclonal antibody-producing hybridomas or T-cell hybridomas¹⁴⁻¹⁶. We have shown here that combining polyethylene glycol with poly-L-arginine makes possible with high efficiency the fusion of

Table 1 Restoration of mitogenic response in mytomycin C-treated C3H/HeN (HeN) spleen cells after fusion with karyoplasts obtained from C3H/HeJ (HeJ) spleen cells

Responding cells	Mitogen		
	None	LPS (50 µg ml ⁻¹)	Con A (5 µg ml ⁻¹)
C3H/HeN	5,653 ± 1,122*	98,570 ± 12,682	34,193 ± 7,563
C3H/HeJ	6,025 ± 1,002	4,796 ± 310	93,029 ± 9,694
MMC-C3H/HeN	320 ± 135	1,804 ± 85	60 ± 3
MMC-C3H/HeJ	337 ± 196	254 ± 48	107 ± 38
C3H/HeJ karyoplasts	193 ± 28	218 ± 28	185 ± 11
Hybrid cells			
MMC-HeN + HeJ karyoplasts	6,586 ± 1,232	72,749 ± 8,738	57,164 ± 5,267
MMC-HeJ + HeJ karyoplasts	944 ± 152	737 ± 111	54,828 ± 20,383

Mouse spleen cells (1×10^8 cells, viability >90%) treated with ammonium chloride to remove erythrocytes were suspended in 3 ml of Eagle's minimal essential medium (MEM) and layered on Ficoll-400 (Pharmacia) gradients (discontinuous gradients of 12.5, 15, 16, 17 and 25%) in Eagle's MEM containing $10 \mu\text{g ml}^{-1}$ of cytochalasin B (Aldrich Biochemicals) and 0.5% dimethyl sulphoxide (Sigma) in cellulose nitrate tubes. Ficoll gradients were sterilized by autoclaving. Cells were centrifuged for 70 min at 23,000 r.p.m. at 31 °C with a Beckman SW-27 swinging bucket rotor in L5-65B ultracentrifuge. After centrifugation, cells in the 17–25% Ficoll fractions were collected and washed twice with 10 vol of MEM by centrifugation at 2,000 r.p.m. for 5 min. More than 95% of the cells in this fraction were karyoplasts and these were further purified by 1g velocity sedimentation using bovine serum albumin gradients (0.5–2% in phosphate-buffered saline). After fractionation with 1g sedimentation, highly purified karyoplasts were obtained (>99% of cells were karyoplasts). Spleen cells were treated with $40 \mu\text{g ml}^{-1}$ of mitomycin C (MMC) for 40 min at 37 °C before cell fusion or culture. MMC-treated spleen cells ($1 \times 10^6 \text{ ml}^{-1}$) were mixed with karyoplasts ($1 \times 10^7 \text{ ml}^{-1}$) in Eagle's MEM. The mixtures were centrifuged for 5 min at 1,500 r.p.m. in conical tubes (Falcon no. 2095), the supernatant thoroughly removed, and three drops of 50% polyethylene glycol (Koch-light MW 6,000) in Eagle's MEM containing 10% dimethyl sulphoxide and $5 \mu\text{g ml}^{-1}$ of poly-L-arginine (Sigma) were added to the cell pellet. The tubes were gently rotated for 1 min at room temperature and the reaction stopped by adding 10 ml of pre-warmed Eagle's MEM. Cells were washed with MEM at room temperature and suspended in RPMI-1640 medium containing 10% fetal calf serum and 2 mM L-glutamine. Spleen cells (2×10^5 per well), karyoplasts (1×10^6 per well) or cells containing the hybrids (2×10^5 per well) were cultured in the presence or absence of mitogen for 72 h and pulsed with $0.5 \mu\text{Ci}$ of ³H-thymidine for the last 4 h. Bacterial LPS was *Salmonella typhosa* W0901 (Difco); Con A was from Sigma. Simple mixtures of MMC-treated HeN spleen cells and HeJ karyoplasts showed only marginal responses to LPS or Con A. MMC-treated HeN cells treated with polyethylene glycol in the same manner did not respond to the above mitogen (data not shown).

* Mean ± s.d. (c.p.m.).

normal lymphocytes with karyoplasts. This simple method for transferring intact nuclei from lymphocytes to inactivated recipient cells could be useful for studying subcellular events in lymphoid differentiation and contribute to our understanding of certain immunological disorders.

We acknowledge the suggestion of Dr T. Sekiguchi to use poly-L-arginine for cell fusion and helpful discussions with Dr Paul W. Kincade. This work was supported in part by grants from the Ministry of Education, Science and Culture, and from the Ministry of Health and Welfare, Japan.

Received 20 October; accepted 16 December 1980.

1. Sulzer, B. M. *Infect. Immun.* **5**, 107–113 (1972).
2. Watson, J. & Riblet, R. *J. exp. Med.* **140**, 1147–1161 (1974).
3. Coutinho, A., Moller, G. & Gronowicz, E. *J. exp. Med.* **142**, 253–260 (1975).
4. Glode, L. M., Sher, I., Osborne, B. & Rosenstreich, D. L. *J. Immun.* **116**, 454–461 (1976).
5. Watson, J., Riblet, R. & Taylor, B. A. *J. Immun.* **118**, 2088–2093 (1977).
6. Forni, L. & Coutinho, A. *Eur. J. Immun.* **8**, 56–62 (1978).
7. Coutinho, A., Forni, L. & Watanabe, T. *Eur. J. Immun.* **8**, 63–67 (1978).
8. Gregory, S. H., Zimmerman, D. H. & Kern, M. J. *Immun.* **125**, 102–107 (1980).
9. Ohara, J., Sekiguchi, T. & Watanabe, T. (in preparation).
10. Wigler, M. H. & Weinstein, I. B. *Biochem. biophys. Res. Commun.* **63**, 669–674 (1975).
11. Watson, J., Kelly, K., Largen, M. & Taylor, B. A. *J. Immun.* **120**, 422–424 (1978).
12. Carter, S. B. *Nature* **213**, 261–264 (1967).
13. Pontecorvo, G. *Somat. Cell Genet.* **1**, 394–400 (1975).
14. Kohler, G. & Milstein, C. *Nature* **256**, 495–497 (1975).
15. Melchers, F., Potter, M. & Warner, N. L. (eds) *Curr. Topics Microbiol. Immun.* **81** (1978).
16. Kennett, R. H., McKearn, T. J. & Bechtol, K. B. (eds) *Monoclonal Antibodies* (Plenum, New York, 1980).

B-cell activation by helper cells is a two-step process

Carlos Martinez-A

Basel Institute for Immunology, Grenzacherstrasse 487,
4005 Basel, Switzerland

Antonio Coutinho

Department of Immunology, Umeå University, 901 85 Umeå,
Sweden

Recent methods for assaying helper activity and cooperative induction of B cells allow for the specific interactions of large populations of lymphocytes in culture^{1,2}. These assays, which do not select for combining-site specificity of the responding B cells, were used to investigate the precise roles of direct cellular recognition and soluble mediators in the process by which specific T-helper cells stimulate resting B cells into clonal expansion and maturation to immunoglobulin secretion. We report here results demonstrating that helper cells perform two distinct functions in the process of B-cell activation: (1) on direct recognition of antigen (and restricting elements) on B-cell surfaces, helper cells induce, in a short time, target B-cell reactivity to soluble growth factors; (2) on functional recognition of antigen on 'stimulator' cells, helper cells also participate in the production of nonspecific growth factors competent to induce proliferation and maturation in those B cells that have been 'activated' by direct interaction. Either of these two steps alone is not sufficient to induce proliferation and maturation of resting B lymphocytes.

Co-culture of helper T cells specific for B-cell surface antigens with 'target' spleen cells leads to activation of large numbers of B cells that can be detected in a polyclonal plaque assay¹. Antigenic spleen cells function in this assay as 'stimulators' of helper cell activation and, in addition, the B-cell compartment of such stimulators initiates clonal expansion and maturation. As the stimulator function of spleen cells, in contrast to the B-cell responses, is radiation resistant², 'bystander' experiments can be performed to determine whether or not B-cell activation by helper cells requires direct cell-to-cell contact. Figure 1 shows the results of an experiment demonstrating that bystander, resting B cells are not activated. C3H/HeJ helper cells specific for 'minor' antigens of the C3H/Tif substrain¹ activate C3H B cells very efficiently but fail to respond to, or to activate, spleen cells of the H-2-compatible B10.Br strain. When the two types of target cells are mixed in the same culture, in which helper-cell responses are ensured by irradiated stimulator cells³, B10.Br B cells are not activated at any effector helper/target B-cell ratio. From this finding, which has also been obtained in a completely syngeneic system using hapten-specific helper cells⁴, we conclude that successful collaboration requires direct T-B-cell interactions.

The apparent requirement for direct cell interactions in cooperative B-cell activation was difficult to reconcile with the many reports that soluble helper factors are released on T-cell activation, as these would be expected to be generated *in situ* by T-cell recognition of the appropriate stimulator cells. The above results indicate that such factors are not competent to initiate the response of resting B cells.

Schreier *et al.* have recently reported that media conditioned by activated helper cells can maintain mitogen-derived blasts, but not resting B cells, in exponential growth⁵. To assess the generation of such B-cell growth factors in our cultures, we repeated the above experiments using lipopolysaccharide (LPS)-activated B-cell blasts as bystander cells. As shown in Fig. 2, activation of helper cells by the irradiated stimulator cells

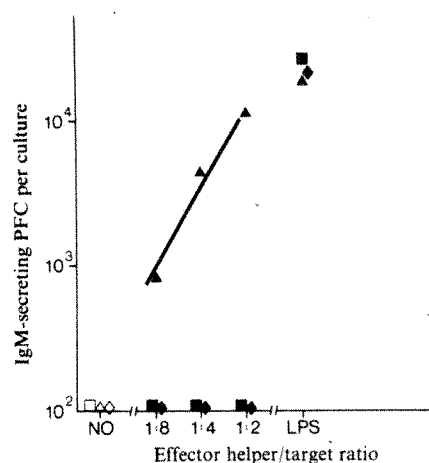


Fig. 1 The helper activity of C3H/HeJ anti-C3H/Tif helper cells, prepared as described elsewhere¹ by *in vivo* priming and *in vitro* enrichment for 4 weeks, was tested by mixing the indicated numbers with 3×10^4 target C3H/Tif (▲) spleen cells in 0.2-ml cultures in RPMI-1640 medium supplemented with 2-mercaptoethanol and 10% fetal calf serum¹. In parallel cultures helper cells were titrated into mixtures of 3×10^4 bystander B10.Br small spleen cells purified by density gradient centrifugation, with the same numbers of irradiated (3,300 rad) C3H/Tif spleen cells (◆). The IgM-secreting plaque-forming cell (PFC) responses were determined in the protein A plaque assay²⁴ 5 days later. Control responses of the same spleen cells to the B-cell mitogen LPS are also shown. Helper cells incubated with resting B10.Br spleen cells alone induce no response (■)¹.

does result in the generation of factors that maintain exponential growth and maturation in unrelated B-cell blasts.

Taken together, these results suggest that helper T cells, in addition to their role in the production of soluble nonspecific growth factors, also induce in resting B lymphocytes a state of responsiveness to those growth factors. This first step in the induction process could presumably explain the requirement for direct interactions. However, this model could only be confirmed by showing that the initial contact of resting B cells with specific helper cells could induce the 'transition' in the reactivity of the responding B cells to growth factors. In fact, the experiments with mitogen-derived B-cell blasts give no information on the initiation of helper cell-dependent induction.

To investigate these possibilities, trinitrophenyl (TNP)-derivatized BALB/c spleen cells were incubated for 48 h with specific BALB/c anti-TNP-BALB helper cells². These cultures were then depleted of T cells and the B-cell blasts originating from this initial T-B-cell interaction were purified and added as bystander cells to cultures where B10.Br anti-NP-B10.Br helper cells were stimulated by the appropriate (NP-self) irradiated spleen cells. As shown in Fig. 3, B-cell blasts activated by the cooperative interaction maintain exponential growth and maturation if unrelated T-helper cell activity takes place in the same culture. Bystander blast responses, however, strictly require concomitant, though unrelated, activation of helper cells, as there is no growth in cultures where all these cell populations were also present but the irradiated B10.Br stimulator cells were not hapten-derivatized. This control excludes 'background' activation of blasts, as well as direct interaction of NP-specific B10.Br helper cells with the BALB/c B-cell blasts. Our results demonstrate that initial T-B-cell interactions result in the acquisition of responsiveness to soluble nonspecific growth factors by the activated B cells. In addition, they show that B cells 'poised' to respond require additional soluble factors to initiate and/or maintain exponential growth.

Cooperative B-cell responses can therefore be separated into two distinct phases of which only the initiation requires direct contact with specific helper cells, whereas expansion is mediated

by soluble B-cell growth factors generated independently by helper cell interaction with competent antigen-presenting cells. It remains to be investigated whether the same helper cell is competent at the two different response levels, or whether functionally distinct T-cell subsets are involved in the initiation and expansion of B-cell responses. The cellular origin of B-cell growth factors, and their molecular characteristics are also unknown. On the other hand, previous experiments in this system have confirmed, with one exception³, the strict major histocompatibility complex restriction of the direct helper cell-B cell interactions⁴, regardless of the nature of the B-cell antigens recognized. It seems, therefore, that I-A- or I-E (C)-encoded⁶ structures on B-cell surfaces are the receptors for the helper cell signals resulting in the transition to responsiveness to B-cell growth factor. As the present experiments were performed in the absence of antigen recognition by the responding B cells whereas activation required specific antigen recognition by helper cells on B-cell surfaces, we suggest that the only role of antigen recognition (immunoglobulin-receptor binding) by B cells is that of making the antigen available on specific B-cell surfaces for T-helper cell recognition, thereby ensuring direct interaction and Ia-mediated induction⁷. These results agree with those demonstrating the necessity for specific helper cells and the failure to substitute for them with soluble factors in the response to protein antigens^{8,9}.

Classical observations, such as the requirements for 'linked recognition' of hapten-carrier conjugates and for the presence of more than one antigenic determinant in immunogenic molecules^{10,11} as well as the unidirectionality of the allogeneic effect *in vivo*^{12,13} can also be explained by this model. More recent results demonstrating the I restriction of T-B-cell interactions^{14,15} and the tendency to 'monogamy' of specific helper cells¹⁶ also fit in with the present schematic framework.

In contrast, the classical bystander experiments of Hartmann¹⁷, Schimpl and Wecker¹⁸ and Dutton *et al.*¹⁹ are in apparent contradiction with the present model. As pointed out by others^{8,9}, however, the *in vitro* responses to sheep erythrocytes are the exception rather than the rule in their lack of requirement for specific helper cells. *In vivo* experiments show no bystander effects and the responses require direct recognition of antigen by helper cells on B-cell surfaces^{15,20}. The *in vitro*

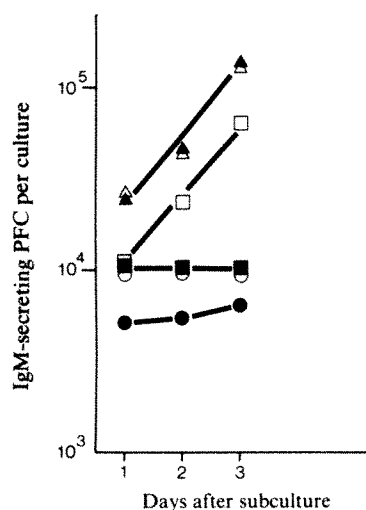


Fig. 2 B10.Br in B-cell blasts, purified by density gradient centrifugation²⁵ after 48 h of LPS stimulation were recultured (3×10^4 per culture) either alone (filled symbols) or with 3×10^4 irradiated (3,300 rad) C3H/Tif spleen cells (open symbols). Both types of cultures were stimulated with either LPS (Δ , $\blacktriangle$) 10^4 C3H/HeJ anti-C3H/Tif helper cells ($\square$, $\blacksquare$), prepared as described in Fig. 1 legend, or left untreated ($\circ$, $\bullet$). The IgM PFC responses were determined at the indicated times.

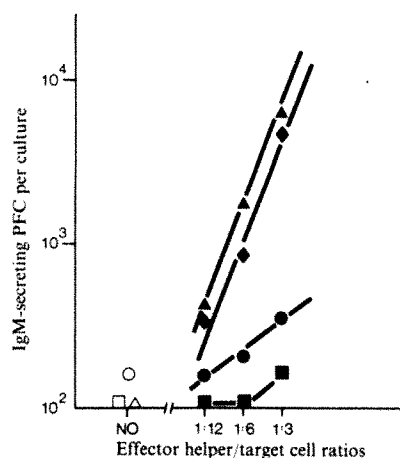


Fig. 3 Specific BALB/c anti-TNP BALB/c and B10.Br anti-NP B10.Br helper cells used in these experiments were obtained as summarized in Fig. 1 legend, and previously described in detail, by immunization with BALB/c spleen cells derivatized with TNP or B10.Br spleen cells coupled with NP², respectively. In a first culture, TNP-derivatized BALB/c small spleen cells, purified by density gradient centrifugation, were incubated with BALB/c anti-TNP BALB/c-specific helper cells at a 3:1 ratio. After 48 h, B-cell blasts were isolated from this culture by density gradient centrifugation, followed by treatment with anti-Thy-1 antibodies and complement, and used in the test cultures shown. These contained B10.Br anti-B10.Br-NP helper cells that were cultured at the indicated concentrations with 3×10^4 non-irradiated, non-derivatized ($\square$) or NP-derivatized ($\blacktriangle$) B10.Br spleen cells. Parallel cultures of helper cells were stimulated with the same number of irradiated (3,300 rad) B10.Br spleen cells, either normal ($\circ$) or NP-derivatized ($\bullet$) and also received 3×10^4 purified BALB/c B-cell blasts isolated from the first culture. The IgM PFC responses were measured on day 4 of culture. Open symbols represent numbers of IgM PFC observed in each type of spleen cell culture, in the absence of helper cells. No PFC responses were observed in mixed cultures of helper cells with irradiated B10.Br spleen cells, either normal or NP-derivatized (not shown).

bystander effects can be made compatible with the present model by attributing the exceptional behaviour of sheep erythrocytes to their ability to induce, in resting B cells, the transition in reactivity to growth factors²¹⁻²³.

We acknowledge the technical assistance of Ms O. Nikkanen and M. Tuneskog and the support of the Swedish MRC.

Received 1 August; accepted 31 December 1980.

1. Augustin, A. A. & Coutinho, A. *J. exp. Med.* **151**, 587-601 (1980).
2. Martinez-A., C. *et al. Eur. J. Immun.* **10**, 403-410 (1980).
3. Coutinho, A. & Augustin, A. A. *Eur. J. Immun.* **10**, 535-541 (1980).
4. Martinez-A., C. *et al. Eur. J. Immun.* **10**, 411-414 (1980).
5. Schreier, M., Andersson, J., Lernhardt, W. & Melchers, F. *J. exp. Med.* **151**, 194-203 (1980).
6. Martinez-A., C., Coutinho, A. & Bernabé, R. R. *Immunogenetics* **10**, 299-303 (1980).
7. Coutinho, A. & Möller, G. *Scand. J. Immun.* **3**, 133-146 (1974).
8. Harwell, L., Kappler, J. & Marrack, P. *J. Immun.* **117**, 1953 (1976).
9. Waldman, H. & Munro, A. J. *J. Immun.* **27**, 53 (1974).
10. Rajewsky, K., Schirrmacher, V., Nase, S. & Jerne, N. K. *J. exp. Med.* **129**, 1131 (1969).
11. Mitchinson, N. A. *Eur. J. Immun.* **1**, 18 (1971).
12. Katz, D. H., Paul, W. E. & Benacerraf, B. *J. Immun.* **107**, 1319 (1971).
13. Rajewsky, K., Roelants, G. E. & Askonas, B. A. *Eur. J. Immun.* **2**, 592-598 (1972).
14. Katz, D., Hamaoka, T., Dorf, M. E. & Benacerraf, B. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2524 (1973).
15. Sprent, J. *J. exp. Med.* **147**, 1159 (1978).
16. Waldmann, H. & Phillips, J. *Springer Semin. Immunopath.* **3**, 129-144 (1980).
17. Hartman, K. Q. *J. exp. Med.* **133**, 1325-1333 (1971).
18. Schimpl, A. & Wecker, E. *Nature* **237**, 15 (1972).
19. Dutton, R. W. *et al. Prog. Immun.* **1**, 355 (1971).
20. Tees, R. & Schreier, M. H. *Nature* **283**, 780-781 (1980).
21. Byrd, W., Feldman, M. & Palmer, J. *Immunology* **27**, 331-337 (1974).
22. Metcalf, D. *et al. J. exp. med.* **142**, 1534-1549 (1975).
23. Claesson, M. H., Layton, J. E. & Luckenback, G. *Immunology* **35**, 397-406 (1978).
24. Gronowicz, E., Coutinho, A. & Melchers, F. *Eur. J. Immun.* **6**, 558-590 (1970).
25. Kurnick, J. T. *et al. Scand. J. Immun.* **10**, 563-573 (1979).

Clones of helper T cells mediate antigen-specific, H-2-restricted DTH

A. T. J. Bianchi, H. Hooijkaas & R. Benner

Department of Cell Biology and Genetics, Erasmus University,
PO Box 1738, 3000 DR Rotterdam, The Netherlands

R. Tees, A. A. Nordin* & M. H. Schreier

Basel Institute for Immunology, Postfach,
CH-4005 Basel 5, Switzerland

It is now well established that in the mouse, helper T cells and killer T cells are two distinct thymus-derived lymphocyte subpopulations, differing from each other in Lyt phenotype and H-2 restriction, among other parameters. Helper T cells are Lyt-1⁺ and their action in immune responses involves restriction at the H-2I region of the major histocompatibility complex (MHC)^{1,2}. Killer cells, on the other hand, are Lyt-23⁺ and their activity is restricted by H-2K/D (refs 1, 3). In most instances, T cells mediating delayed-type hypersensitivity (DTH) responses share the Lyt phenotype and H-2 restriction of the helper T cell⁴⁻⁷. This raises the question of whether or not helper activity and DTH can be mediated by the same activated T cell. Arguments for both views have been reported⁸⁻¹². We analysed this question using clones of specific helper T cells, which were obtained by long-term culture *in vitro* of *in vivo* primed T cells, followed by single-cell cloning¹³⁻¹⁵. Here we show that these clones of helper T cells mediate antigen-specific and fully H-2-restricted DTH. The restricting element lies to the left of the I-B region in genetic maps of the mouse MHC.

Long-term cultured, cloned and subcloned sheep or horse red blood cell-specific helper T cells (T'_{SRC}, T'_{HRC} respectively), which have been maintained for 5-24 months in culture and previously shown to function as helper T cells *in vitro*¹³⁻¹⁵ and *in vivo*¹⁵, were examined for DTH activity. This DTH responsiveness was assayed by means of the immune lymphocyte transfer assay¹⁶. As few as 1,000 cloned helper T cells, injected together with SRC into syngeneic recipients, caused a significant DTH response which increased in severity as the dose of T cells increased (Fig. 1). In comparison with *in vivo* activated lymph node cells, the cloned helper T cells were far more effective on a per cell basis. At least 10³ times more *in vivo* activated cells were required to induce a response of similar magnitude (Fig. 1, right). As controls, the other footpad of the mice was injected with cloned helper T cells without antigen or with an irrelevant antigen—in neither case was a DTH response induced. DTH responses were mediated equally well by other clones of T'_{SRC} and T'_{HRC} and by four different clones of egg albumin-specific helper T cells when injected together with the relevant antigen.

The H-2 restriction of cellular interactions involved in DTH responses¹⁷ was examined using a clone of T'_{SRC} and a subclone of T'_{HRC}. Absolute H-2 restriction was observed in both instances. Results obtained with T'_{SRC} (clone 26-14) are shown in Fig. 2. The C57BL/6J-derived cloned helper T cells mediated DTH responses only in the C57BL/6J and B10.A(5R) recipients. B10.HTT mice (incompatible with C57BL/6J mice for the whole H-2 complex) and B10.A(4R) mice (incompatible at H-2K and I-A) did not show any DTH reactivity on injection of the cloned helper T cells plus antigen. It should be stressed that this result was found after transfer of 3.3 × 10⁴ viable T cells, which is about 30 times the number of cells required to induce a significant response in a syngeneic combination. This result shows that the T cell-macrophage interactions which underlie

these DTH responses¹⁷ are restricted by either H-2K or I-A. Appropriate congenic strains were not available at the time of these experiments for distinguishing between K- or I-A-region restriction. However, the action of helper T cells in immune responses *in vivo* has been demonstrated to involve restriction only at I-A².

The antigen-specific and H-2-restricted induction of DTH response with cloned helper T cells in normal mice does not exclude the participation of host T cells at the effector level. We have therefore tested the DTH response induced by cloned helper T cells in congenic *nu/nu* mice (Fig. 2). The DTH response in these T cell-deficient animals showed the same kinetics as and was of similar magnitude to that observed in thymus-bearing, I-A-compatible mice. This result indicates that host T cells are not essential for the response and that the cloned helper T cells are the DTH effector cells.

Many investigations have attempted to establish the relationship between helper T cells and DTH effector cells⁸⁻¹². Although some studies pointed to an inverse relationship between the *in vivo* occurrence of activated helper T cells and DTH effector cells⁸⁻¹⁰, others have found that helper activity and DTH can coincide^{11,12}. These experiments were carried out with uncloned T-cell populations and failed to prove that both responses can be mediated by the same cell. However, our experiments show unequivocally that helper T cells can mediate DTH responses which were observed not only in the local immune lymphocyte transfer assay, but also after intravenous injection of these cells. Mice infused with 1-3 × 10⁶ cloned helper T cells responded to antigen challenge to the hind footpad with a smaller but significant DTH response. This requirement for much higher numbers of cells may be due to changes in recirculation and

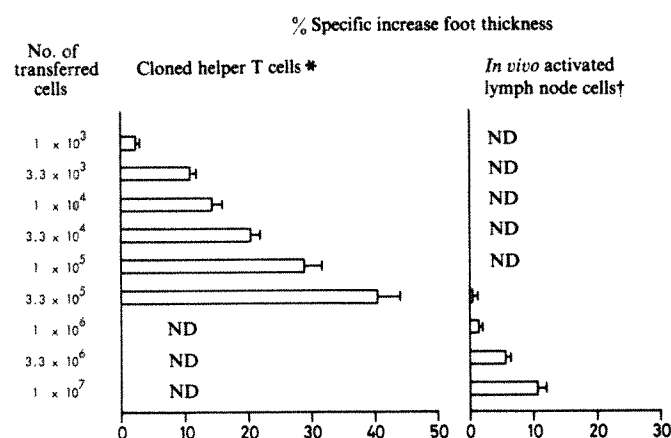
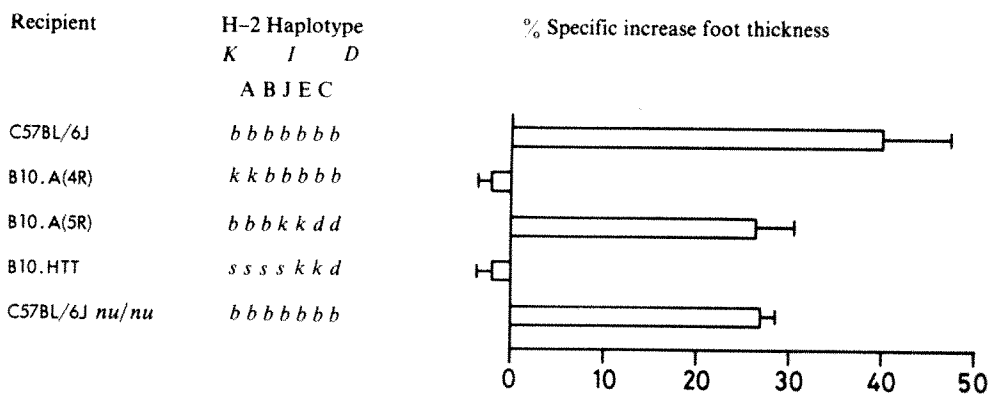


Fig. 1 DTH sensitivity by different numbers of long-term cultured C57BL/6J-derived, SRC-specific helper T cells (clone S26-14) (left), and by different numbers of lymph node cells which were activated *in vivo* by subcutaneous immunization of C57BL/6J mice with 2 × 10⁷ SRC (right). The cells to be tested were injected together with 5 × 10⁷ SRC in a volume of 50 µl into the instep of the right hind leg of C57BL/6J female mice (Institut für Biologisch-Medizinische Forschung AG). Recipient mice injected with cloned SRC-specific helper T cells plus SRC in the right hind leg received the same number of HRC-specific helper T cells (subclone 18-33-3) together with 5 × 10⁷ SRC in the left hind leg. Recipient mice which received *in vivo* activated lymph node cells plus SRC in the right hind leg received the same number of non-activated lymph node cells and 5 × 10⁷ SRC in the left hind leg. At 24 h after injection the thickness of the hind feet was measured. DTH responses are expressed as the specific % increase in foot thickness, and were calculated as (thickness right foot - thickness left foot)/thickness left foot × 100%. Horizontal bars represent 1 s.e.m. (n = 7). ND, not determined. *, The establishment and maintenance of the long-term cultured specific helper T cells has been described previously¹³⁻¹⁵. †, The activated lymph node cells were obtained by collecting the inguinal and axillary lymph nodes of C57BL/6J mice 4 days after subcutaneous immunization with 2 × 10⁷ SRC, equally distributed over the inguinal and axillary areas. The non-activated lymph node cells were obtained from naive C57BL/6J mice.

* On leave of absence from the NIH, NIA, Bethesda, Maryland 20205, USA.

Fig. 2 DTH reactivity mediated by cloned SRC-specific helper T cells was tested for H-2 restriction. Recipient mice (all from the Institut für Biologisch-Medizinische Forschung AG) were injected with 3.3×10^4 SRC-specific helper T cells (clone S26-14) and 5×10^7 SRC in a volume of $50 \mu\text{l}$ in the instep of the right hind foot, and with 3.3×10^4 HRC-specific helper T cells (subclone 18-33-3) and 5×10^7 SRC in the instep of the left hind foot. At 24 h after transfer the thickness of the hind feet was measured. DTH responses are expressed as the specific percentage increase in foot thickness and calculated as described in Fig. 1 legend. Horizontal bars represent 1 s.e.m. ($n = 7$).



homing properties brought about by long-term *in vitro* propagation. Similar observations were made in clonally reconstituted *nu/nu* mice where $1-5 \times 10^6$ of the same cloned helper T cells were required to induce specific antibody responses comparable with or higher than those observed in thymus-bearing mice¹⁵.

Obviously, our results do not argue against the possibility that T cells other than helper T cells may also mediate DTH reactivity. Several studies have shown that DTH to allogeneic cells can be transferred by Ly-2⁺ cells⁷ and that contact sensitivity to dinitrofluorobenzene⁶ and DTH to non-H-2 alloantigens¹⁸ and viral antigens¹⁹ can be H-2K/D restricted. Indeed, it has been reported that T cells mediating DTH and cytotoxic T cells could not be separated by size, density or antigen-binding properties²⁰.

The clones of helper T cells used for these studies have been shown to release several biologically active mediators, the induction of which is strictly antigen specific and H-2 restricted^{13,14,21-23}. So far, factors affecting B-cell^{21,22} and T-cell¹⁴ growth and *in vitro* colony formation by granulocytic, macrophage and erythroid precursor cells²³ have been discerned. Soluble factors have been indicated in the induction of the DTH lesion²⁴. Histologically, the DTH response mediated by cloned helper T cells in *nu/nu* and normal mice was indistinguishable from other Jones-Mote-type DTH responses. The nature of the cellular infiltration strongly indicates the release of chemotactic factors. Clones of helper T cells may help to characterize these activities. In addition, the DTH response is the most rapid assay to date for screening cloned and uncloned helper T cells for specificity and H-2 restricted activity. The assay is as sensitive and reproducible as the *in vitro* assay for specific helper activity. Moreover, the cellular nature of the I-A-bearing, antigen-presenting cells may be identified by their simultaneous injection with helper T cells into I-A-incompatible hosts.

We thank Drs K. Fischer-Lindahl and H. v. Boehmer for their critical reading of the manuscript. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-LaRoche and Co. Ltd.

Received 25 November 1980; accepted 13 January 1981.

- Cantor, H. & Boyse, E. A. *J. exp. Med.* **141**, 1376-1389 (1975).
- Sprent, J. *Immun. Rev.* **42**, 108-148 (1978).
- Bevan, M. J. *J. exp. Med.* **142**, 1349-1364 (1975).
- Miller, J. F. A. P., Vadas, M. A., Whitelaw, A. & Gamble, J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5095-5098 (1975).
- Vadas, M. A. *et al. J. exp. Med.* **144**, 10-19 (1976).
- Vadas, M. A., Miller, J. F. A. P., Whitelaw, A. M. & Gamble, J. R. *Immunogenetics* **144**, 137-153 (1977).
- Smith, F. & Miller, J. F. A. P. *Int. Archs Allergy appl. Immun.* **58**, 295-301 (1979).
- Parish, C. R. in *Progress in Immunology* Vol. 2 (eds Brent, L. & Holborow, J.) 39-44 (North-Holland, Amsterdam, 1974).
- Lubet, M. T. & Kettman, J. R. *J. Immun.* **123**, 426-433 (1979).
- Ramshaw, J. A. & Eidenger, D. *Cell. Immun.* **42**, 42-47 (1979).

- Lagrange, P. H. & Mackaness, G. B. *J. exp. Med.* **148**, 235-245 (1978).
- Kerckhaert, J. A. M. & Hofhuis, F. M. A. *Annls Inst. Pasteur, Paris* **126C**, 371-382 (1975).
- Schreier, M. H. & Tees, R. *Int. Archs Allergy appl. Immun.* **61**, 227-237 (1980).
- Schreier, M. H., Iscove, N. N., Tees, R., Aarden, L. & von Boehmer, H. *Immun. Rev.* **51**, 315-336 (1980).
- Tees, R. & Schreier, M. H. *Nature* **283**, 780-781 (1980).
- Brent, L., Brown, J. B. & Medawar, P. B. *Proc. R. Soc. B* **156**, 187-209 (1962).
- Lubet, M. T. & Kettman, J. R. *Cell. Immun.* **45**, 151-161 (1979).
- van der Kwast, T. H. *J. Immunogenet.* **7**, 315-324 (1980).
- Zinkernagel, R. M. *J. exp. Med.* **144**, 776-787 (1976).
- Elliott, B. E., Haskill, J. S. & Axelrad, M. A. *J. exp. Med.* **141**, 584-599 (1975).
- Schreier, M. H., Andersson, J., Lernhardt, W. & Melchers, F. *J. exp. Med.* **151**, 194-203 (1980).
- Melchers, F., Andersson, J., Lernhardt, W. & Schreier, M. H. *Eur. J. Immun.* **10**, 679-685 (1980).
- Schreier, M. H. & Iscove, N. N. *Nature* **287**, 228-230 (1980).
- Phillips, S. M., Carpenter, C. B. & Merrill, J. P. *Cell. Immun.* **5**, 249-263 (1972).

Altered expression of the calcitonin gene associated with RNA polymorphism

Michael G. Rosenfeld*, Susan G. Amara*,
Bernard A. Roos†, Estelita S. Ong‡
& Ronald M. Evans‡

* Division of Endocrinology, School of Medicine, University of California, San Diego, La Jolla, California 92093, USA
† Endocrinology and Mineral Metabolism, Veterans Administration Medical Center and School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA
‡ Tumor Virology Laboratory, The Salk Institute, San Diego, California 92138, USA

The selective control of gene expression results in diversified morphology and physiological function. Understanding the expression of differentiated function in molecular terms requires detailed characterization of the regulation of mRNA synthesis and catabolism. Although considerable emphasis has been placed on transcriptional control, the discovery of HnRNA and 'split genes' gives rise to the possibility of post-transcriptional regulation at the level of processing of nuclear precursors^{1,2}. A rat calcitonin-producing medullary thyroid carcinoma (MTC) line was used as a model for definition of certain aspects in regulation of gene expression. Serial transplantation of several rat MTC lines containing and secreting large amounts of calcitonin generated tumours in which calcitonin biosynthesis was decreased more than 10-fold. We report here that the conversion from a 'high' to a 'low' calcitonin producing state is associated with specific modifications of the calcitonin mRNA synthetic pathway and a consequence of these changes seems to be the production of a new cytoplasmic mRNA.

Variation in calcitonin biosynthesis typifies a frequent event in many hormone-producing tumours characterized by a spontaneous reduction in synthesis of a particular hormone, often in association with the appearance of new and frequently predictable hormones³. Thyroid 'C'-cell hyperplasia and ultimate development of MTCs occurs in 10–40% of older rats⁴. Production of calcitonin by these tumours may decrease spontaneously during a single transplantation or following reintroduction of primary culture cells, or gradually during serial transplantation, from the usual 1.5–4.5% of total protein to <0.3% of total protein⁵. The direction of change is characteristically from high to low calcitonin production. Immunohistochemical staining with specific calcitonin antisera reveals that ~95% of the cells in high-producer tumours and <5% of cells in low-producer tumours are calcitonin positive. The conversion to low calcitonin production is frequently accompanied by increased production of somatostatin⁶. These tumours also produce other polypeptide hormones⁷.

Characterization of rat calcitonin mRNA has allowed detailed analysis of calcitonin biosynthesis. mRNA from MTCs programmes the cell-free synthesis of a calcitonin precursor peptide of molecular weight (MW) 17,500 (ref. 8) (alternatively estimated as 15,000 (ref. 9)) and a less prevalent product of MW 18,800 (ref. 8). DNA sequence analysis of a chimaeric plasmid containing DNA complementary to calcitonin mRNA establishes that the 32-amino acid calcitonin sequence is encoded in the 3'-proximal portion of the mRNA¹⁰. Using this clone as probe, normal rat thyroid and high-producer rat MTCs

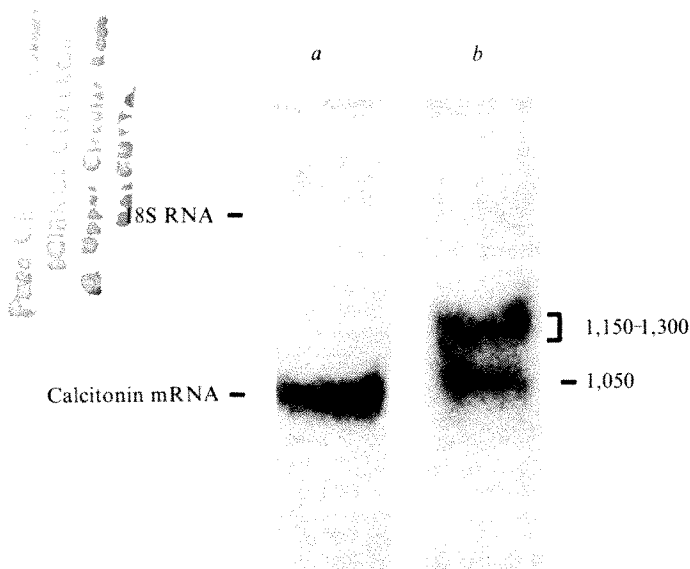


Fig. 1 Analysis of RNA species hybridizing to calcitonin cDNA. Poly(A)-rich RNA was prepared as described elsewhere^{8,11} from high- and low-producer tumours. Aliquots were denatured by reaction with glyoxal¹² and electrophoresed on 1.5% agarose slab gels using 36 mM Tris-HCl (pH 7.6), 30 mM NaH₂PO₄·H₂O, 1 mM EDTA¹³. The gel was stained using acridine orange, photographed and gently shaken for 20 min with 50 mM NaOH, then neutralized with two washes of 0.2 M sodium acetate (pH 4.3). The RNA was then transferred to diazotised paper, washed in prehybridization solution, and hybridized for 14 h at 42 °C against 5×10^5 c.p.m. ml⁻¹ of ³²P-labelled calcitonin cDNA, produced by nick-translation of calcitonin-containing plasmid to a specific activity of 1.5×10^8 c.p.m. μ g⁻¹ (refs 13, 14) using dCTP as the labelled nucleotide. The paper was washed and autoradiographed as previously described. The migration of 18S and 28S ribosomal RNAs and of prolactin mRNA was used to calculate size; the portion of the gel which was blotted contained 18S RNA, as shown. The 1,050-nucleotide species is referred to as calcitonin mRNA. *a*, Poly(A)-rich RNA from a high-producer tumour (10 μ g); *b*, poly(A)-rich RNA from a low-producer tumour (10 μ g).

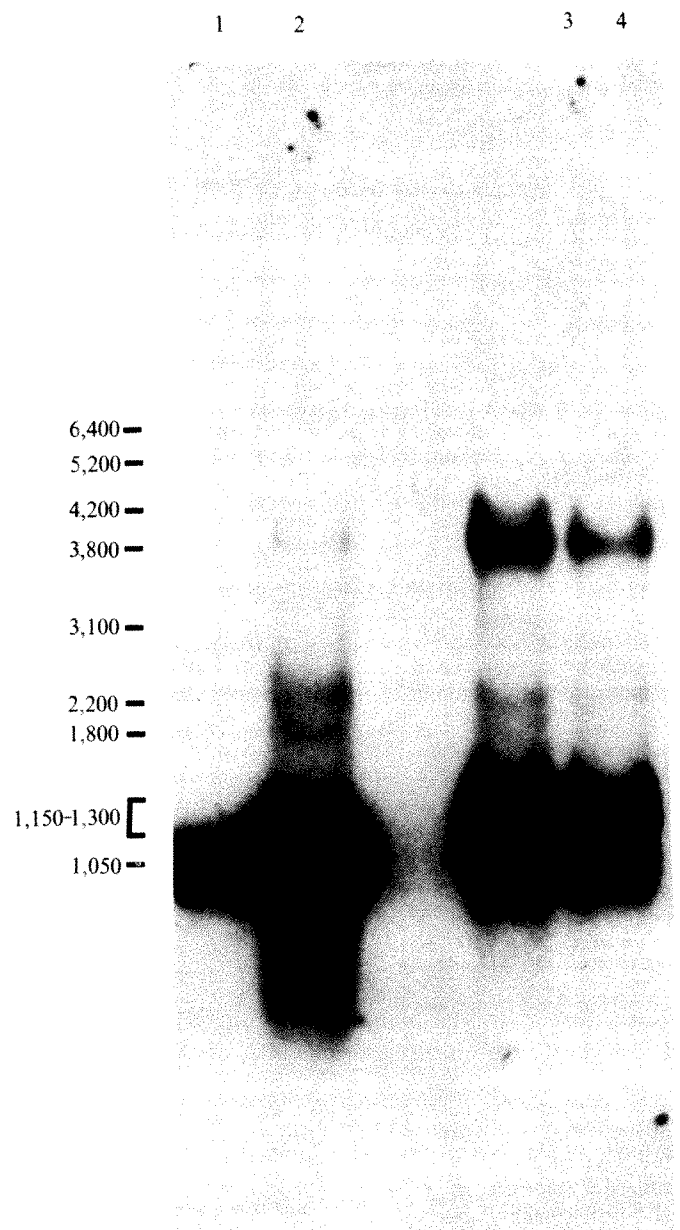


Fig. 2 Analysis of total cellular RNA species hybridizing to calcitonin cDNA. Poly(A)-rich RNA was prepared and processed as described in Fig. 1 legend. The autoradiograph was exposed for 4 days to visualize clearly the larger reactive RNA species. Lane 1, cytoplasmic poly(A)-rich RNA (5 μ g) from a high-producer tumour; lane 2, total poly(A)-rich RNA (10 μ g) from a high-producer tumour; lane 3, total poly(A)-rich RNA (10 μ g) from a low-producer tumour; and lane 4, total poly(A)-rich RNA (10 μ g) from a low-producer tumour of independent origin.

were shown to contain one major poly(A)-rich cytoplasmic RNA species $1,050 \pm 50$ nucleotides in length¹⁰, referred to as calcitonin mRNA.

To identify mechanisms by which variation in calcitonin gene expression occurs, total poly(A)-rich RNA for high- and low-producer tumours was examined. Analyses were carried out by hybridizing labelled, cloned calcitonin cDNA to RNA transferred to diazotised paper after denaturation and size-fractionation through agarose gels (see Fig. 1 legend). In the conditions used, this method provided an accurate and reproducible quantification of RNA species which varied from 1,000 to 5,500 nucleotides in length over a 200-fold range of

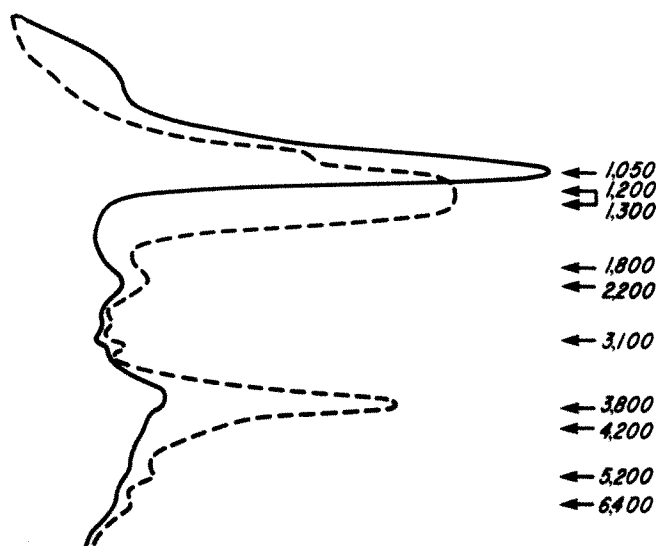


Fig. 3 Scan analysis of autoradiographs of calcitonin cDNA-reactive species. Autoradiographs of calcitonin cDNA-reactive RNA were quantified by scanning with a Helena Quickscan (1-mm window). The sizes of various RNA species are indicated, based on migration of 18S and 28S rRNA. —, High- and - - -, low-producer tumour RNA.

RNA concentration. Remarkable differences in the abundance and size of reactive poly(A)-rich RNA from low-producer tumours were observed. As shown in Fig. 1b, low-producer tumours contain a 1,050-nucleotide transcript which co-migrates with calcitonin mRNA of both high-producer tumours and normal thyroid gland, and a second abundant species which migrates at 150–250 nucleotides longer than calcitonin mRNA (Fig. 1a). Note that this second species seems to lack a similar correspondent in the high-producer lines which suggests that conversion from high to low production is associated with an altered expression of the calcitonin gene or a closely related gene.

To examine this relationship more closely, it is necessary to identify the steps involved in the synthesis of mature calcitonin mRNA. Total poly(A)-rich RNA from high- and low-producer tumours was examined for putative precursor transcripts. In addition to the prominent 1,050-nucleotide cytoplasmic species, RNA from the high-producer tumour contains several calcitonin cDNA-reactive RNA species (Fig. 2, lanes 1 and 2). These include transcripts of 6,400, 5,200, 4,200, 3,800, 3,100, 2,200 and 1,800 nucleotides in length.

Analysis of RNA from two independent low-producer tumours revealed surprising alterations in abundance of most of the nuclear transcripts. There were large increases in the 6,400-, 5,200-, 4,200- and 3,800-nucleotide species whereas the 1,800-nucleotide transcript was significantly decreased (Fig. 2, lanes 3 and 4). The abundant RNA species are overexposed on the autoradiograph (Fig. 2) to visualize the less-abundant putative nuclear RNA species.

Scans of the autoradiographs (Fig. 3) allowed quantification of reactive RNA species. The amount of the 3,800-, 5,200- and 6,400-nucleotide RNA species in low-producer tumours was increased 10- to 30-fold above that in high-producer tumours. In the low-producer tumour, the 3,800-nucleotide RNA species alone bound up to 25–30% of the total probe hybridized. There was no alteration between high- and low-producer tumours in the amount of reactive 3,100- or 2,200-nucleotide RNA species.

Conversion to a low calcitonin producer state was associated with a series of reproducible alterations in calcitonin gene expression. Cells which contained little or no detectable calcitonin transcribe large amounts of calcitonin-reactive nuclear RNA species and generate new cytoplasmic RNA species, 150–250 nucleotides larger than the 'mature' calcitonin mRNA.

The molecular alterations which occur during conversion from high to low calcitonin production could result from one of several possible mechanisms: (1) an increased transcriptional activity of a related gene (or genes), with concomitant reduced expression of the calcitonin gene, corresponding to the reduction in calcitonin synthesis; (2) an alteration in the primary structure of a single calcitonin gene, with resultant alteration in processing of the primary transcript; or (3) the observed RNA heterogeneity in low-producer tumours arises from differential or 'abnormal' processing of one primary RNA transcript. In the latter two cases, the increased abundance of the putative nuclear precursor RNAs could correspond to an altered processing pathway; any of the proposed models could result in the observed RNA polymorphism. To investigate these possibilities, a chimaeric plasmid containing DNA complementary to the additional RNA sequence unique to the new 1,250 RNA present in low-producer tumours has been constructed. This will allow initiation of structural and functional analysis of the new cytoplasmic RNA species, the putative nuclear RNA transcripts, and of the calcitonin gene.

We thank Roger Birnbaum for support during this study and Maureen Gilmore-Hebert and Geoffrey Wahl for their advice and critical review of the manuscript. The work was supported by American Cancer Society grants BC-169 and BC-282A, and by NIH research grants AM18477, AM21567 and GM26444.

Received 1 December; accepted 22 December 1980.

1. Darnell, J. E. *Prog. Nucleic Acid. Res. molec. Biol.* **22**, 327–357 (1978).
2. Gilbert, W. *Nature* **271**, 501 (1978).
3. Baylin, S. B. & Mendelson, G. *Endocrine Rev.* **1**, 45–77 (1980).
4. Boorman, G. A., Hollander, C. F. & Van Noord, M. J. *Archs Path.* **94**, 35–41 (1972).
5. Roos, B. A. *et al. Endocrinology* **105**, 27–32 (1979).
6. Birnbaum, R. S., Muszynski, M. & Roos, B. A. *Cancer Res.* (in the press).
7. Zeytinoglu, F. N., Gagel, R. F., Tashjian, A. H. Jr, Hammer, R. A. & Leeman, S. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3741–3745 (1980).
8. Amara, S. G., Rosenfeld, M. G., Birnbaum, R. S. & Roos, B. A. *J. biol. Chem.* **255**, 2645–2648 (1980).
9. Jacobs, J. W., Potts, J. T. Jr, Bell, N. H. & Habener, J. F. *J. biol. Chem.* **254**, 10600–10603 (1979).
10. Amara, S. G., David, D. N., Rosenfeld, M. G., Roos, R. A. & Evans, R. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4444–4448 (1980).
11. Shields, D. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2059–2063 (1977).
12. McMasters, G. K. & Carmichael, G. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835–4838 (1977).
13. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350–5354 (1977).
14. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).

Dual expression of λ genes in the MOPC-315 plasmacytoma

Alfred L. M. Bothwell, Michael Paskind, Richard C. Schwartz, Gail E. Sonenshein*, Malcolm L. Gefter & David Baltimore

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The expression of two distinct κ light chain immunoglobulins in the MPC-11 mouse myeloma is well established^{1–4}, the two protein products being apparently from RNA transcripts derived from separate, rearranged κ alleles in the MPC-11 genome^{5–7}. Recently, the characterization of κ -related RNAs and protein products in several λ -producing myelomas has indicated that multiple expression of light chain RNAs is a common event in myelomas and other cells of the B-lymphocyte lineage⁸. These studies suggest that, although many light chain alleles may function to make RNA and protein in a given B-lymphocytic cell, only one complete, functional light chain is generally translated from the RNAs present in a single cell. The myeloma, MOPC-315, synthesizes and secretes an antibody

* Present address: Boston University Medical School, Department of Biochemistry, Boston, Massachusetts 02118, USA.

which has an α heavy chain and a λ_{II} light chain⁸. The DNA of MOPC-315 either has no κ genes or has only a fragment of one, but it certainly has no κ genes in the embryonic configuration⁵. Rearrangement of its λ genes has been observed but the exact nature of the rearrangement is not known⁹. Because initial observations suggested that an immunoglobulin-related protein other than α and λ_{II} was present in MOPC-315 cells¹⁰, we undertook to derive molecular cDNA clones from the mRNA in MOPC-315 tumour cells. Analysis of the clones has now identified two λ chain mRNA species: a normal λ_{II} chain mRNA and another which directs the synthesis of a deleted form of a λ_I protein. The nucleotide sequence of the deleted λ_I mRNA shows that it resulted from a joining of the sequence encoding amino acid 31 of the variable region directly to the constant region coding sequence.

To characterize the λ RNA and protein products in MOPC-315 we first derived cDNA clones from total poly(A)-containing RNA of the tumour. In parallel, *in vitro* translation studies on partially purified mRNAs and RNA size analysis with cloned λ probes were performed¹⁰. Clones were derived using the dG-dC homopolymer tailing method to insert DNA in the *Pst*I site of pBR322. After transformation of *Escherichia coli* χ 1776, recombinant plasmid-containing cells were selected that hybridized to a full length λ_I cDNA clone previously derived from myeloma MOPC-104E (A.L.M.B., unpublished). Preliminary restriction enzyme digest patterns indicated the presence of two types of λ cDNA clones. Two of the longest clones representative of the two classes of inserts were chosen for detailed examination. Because one was shown to have λ_{II} sequence and the other λ_I sequence, they were named $p\lambda_{II}$ -1 and $p\lambda_I$ -13.

The relationship of the clones to cellular mRNAs was studied by using the cDNA clones to select mRNA by hybridization and

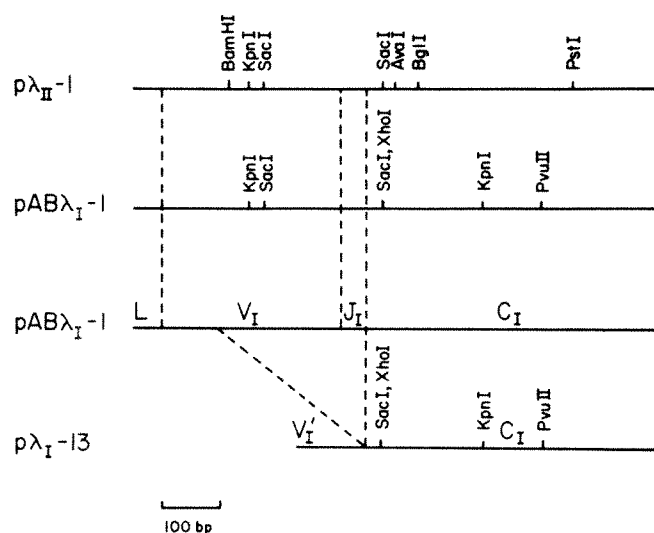


Fig. 2 Restriction maps of the λ_I and λ_{II} cDNA clones. The maps indicate the deletion of the 3'-proximal two-thirds of the variable region and J region from $p\lambda_I$ -13 when compared with the normal λ mRNA. The $pAB\lambda_I$ -1 clone (from MOPC-104E, a λ_I producer) contains a portion of the leader sequence whereas the $p\lambda_I$ -13 clone begins at amino acid 8 in the mature protein. The embryonic λ_{II} variable sequence does not contain a *Bam*HI site¹³. A different amino acid occurs at amino acid position 38 (isoleucine) in the MOPC-315 λ_{II} sequence from that predicted from the germ-line λ_{II} variable gene DNA sequence (valine)¹³. The codon ATC for residue 38 (isoleucine) would generate a *Bam*HI site which is found in $p\lambda_{II}$ -1. bp, base pairs.

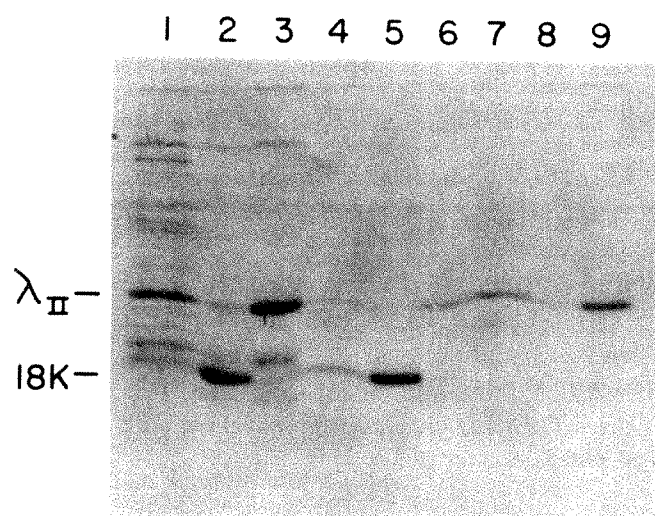


Fig. 1 *In vitro* translation of clone selected mRNA. Either an *Eco*RI-*Xho*I 1,100-base pair DNA fragment derived from $p\lambda_I$ -13 plasmid (10 μ g) or a *Pvu*II-*Ava*I 1,800-base pair DNA fragment derived from 10 μ g of $p\lambda_{II}$ -1 containing the λ_{II} constant region was denatured and bound to a nitrocellulose filter and hybridization and elution procedures performed¹⁸. Membrane-bound poly(A)-containing polysomal RNA (5 μ g) derived from MOPC-315 (tumours maintained by Dr H. Eisen) was hybridized in 65% formamide, 400 mM NaCl, 10 mM PIPES, pH 7.4, at 45 °C for 2 h. The filters were washed 10 times with 150 mM NaCl, 20 mM Tris-HCl pH 7.8, 1 mM EDTA, 0.5% SDS at 65 °C, then three times in 20 mM Tris-HCl pH 7.8, 1 mM EDTA. RNA was removed by boiling for 1 min with 300 μ l of H₂O, ethanol precipitated and translated in a rabbit reticulocyte lysate. Lanes contain total polysomal RNA translation products (lanes 1-3), $p\lambda_I$ -13-selected mRNA translation products (lanes 4-6) or $p\lambda_{II}$ -1-selected mRNA translation products (lanes 7-9). Products in lanes 2, 5 and 8 were precipitated with anti- λ_I serum and those in lanes 3, 6 and 9 with anti- λ_{II} serum. The sources of the antisera and immunoprecipitation procedures are described elsewhere¹⁰.

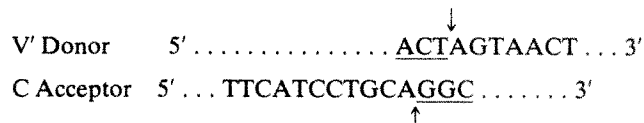
then translating the mRNAs *in vitro* (Fig. 1). Two λ -related protein products could be detected by *in vitro* translation of mRNA. When the total translation products (lane 1) were submitted to specific immunoprecipitation, an 18,000 molecular weight (18K) protein was precipitated with anti- λ_I serum (lane 2) and the complete λ_{II} light chain was precipitated with anti- λ_{II} serum (lane 3). When the mRNA selected by $p\lambda_I$ -13 was translated, an 18K protein was enriched (lane 4) which was precipitable with anti- λ_I serum (lane 5) but not with anti- λ_{II} serum (lane 6). By contrast, when $p\lambda_{II}$ -1 was used to select mRNA, the selected mRNA translated into λ_{II} protein (lane 7) which was not precipitable with anti- λ_I serum (lane 8) but was precipitable with anti- λ_{II} serum (lane 9). We have independently shown that the $p\lambda_{II}$ -1 clone hybridizes to a mature size light chain mRNA (1,100 base pairs) whereas the $p\lambda_I$ -13 clone hybridizes to a mRNA of about 900 base pairs¹⁰. Thus, it is evident that MOPC-315 cells have two types of λ -related mRNA which gave rise to two classes of clone: $p\lambda_{II}$ -1 derived from a full-length mRNA encoding the λ_{II} protein of MOPC-315 antibody and $p\lambda_I$ -13 derived from a truncated λ_I mRNA that translates to form a truncated λ_I -related protein.

The $p\lambda_I$ -13 clone was further analysed by comparing its restriction enzyme digest pattern with that of a clone derived from the full-length λ_I mRNA of myeloma MOPC-104E prepared in this laboratory (A.L.M.B., unpublished) (Fig. 2). The authentic λ_I clone had six identifiable sites, two of which (*Xho*I and *Kpn*I) in the constant region would not have been predicted from the reported 138-base pair sequence¹¹. Sequence data, reported below, have confirmed the existence of these sites in this and in other λ_I cDNA clones; presumably, there were sequencing errors in the previous report. The $p\lambda_I$ -13 clone from MOPC-315 had all the sites found in the authentic λ_I constant region sequence. These distinguish the cloned insert from the λ_{II} genes, the sites of which are quite different from those in the λ_I gene (Fig. 2).

To analyse the structure of $p\lambda_I$ -13 further, we determined the nucleotide sequence of its insert (Fig. 3). The entire constant region sequence was evident but the variable region sequence contained codons that, by comparison with the reported protein

The origin of the λ_1 mRNA in MOPC-315 is uncertain. A DNA rearrangement has probably occurred because the λ_1 variable and constant regions are found on a single mRNA. Analysis of λ genes in MOPC-315 DNA indicates that the only detectable λ_1 rearrangement involving both the variable and constant regions is the normal V_1 to C_1 rearrangement. The 7.4-kilobase fragment co-migrates with the fragment observed in the λ_1 producer, MOPC-104E (A.L.M.B. and D.B., unpublished results). Therefore, the rearrangement, or somatic mutation, must have altered the normal donor site at the 3' end of the J segment. Whether the entire λ_1 variable region is present in the DNA is uncertain and will necessitate fine structure analysis: the deletion of a variable region sequence in the mRNA could result from either an aberrant DNA rearrangement that removed the coding sequence for amino acids 32-109 or from an aberrant RNA splicing event. Precedent for aberrant splicing as a cause of variable region deletion in a mRNA comes from the case of MPC-11 (refs. 6, 7), in which on one chromosome a κ variable

The possibility of an aberrant splice as the cause of the truncated λ_1 mRNA in MOPC-315 is supported by an examination of the nucleotide sequence of the λ_1 variable region. The sequence TAGTAATA that occurs immediately 3' to the deletion site in the complete λ_1 variable region¹¹ has the form close to the canonical donor splice sequence^{14,15}. It is also clear from a comparison of the cleavage points demanded by the sequence of the cDNA clone (indicated by arrows)



This work was supported by grants from the American Cancer Society and the NCI. A.L.M.B. is a postdoctoral fellow of the Medical Foundation. R.C.S. was supported by NIH training grant GMO07287. G.E.S. is an Established Investigator of the American Heart Association. D.B. is a Research Professor of the American Cancer Society.

1. Kuehl, W. M. *et al. Cell* **5**, 139-147 (1975).
2. Burstein, Y., Zeurell, R., Kantor, F. & Schechter, I. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3157-3161 (1977).
3. Rose, S. M., Kuehl, W. M. & Smith, G. P. *Cell* **12**, 453-462 (1977).
4. Perry, R. P., Kelley, D. E. & Schibler, U. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3678-3682 (1979).
5. Alt, F. W., Enea, V., Bothwell, A. L. M. & Baltimore, D. *Cell* **21**, 1-12 (1980).
6. Seidman, J. G. & Leder, P. *Nature* **286**, 279-283 (1980).
7. Choi, E., Kuehl, M. & Walli, R. *Nature* **286**, 279-283 (1980).
8. Dugan, E. S., Bradshaw, R. A., Simms, E. S. & Eisen, H. N. *Biochemistry*, **12**, 5400-5416 (1973).
9. Maki, R., Kearney, J., Paige, C. & Tonegawa, S. *Science* **209**, 1366-1369 (1980).
10. Schwartz, R. C., Sonenshein, G. E., Bothwell, A. & Geffer, M. L. *J. Immun.* (in the press).
11. Bernard, O., Hozumi, M. & Tonegawa, S. *Cell* **15**, 1133-1144 (1978).
12. Appella, E. *Proc. natn. Acad. Sci. U.S.A.* **68**, 590-594 (1971).
13. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1485-1489 (1978).
14. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853-4857 (1978).
15. Catterall, J. F. *et al. Nature* **275**, 510-513 (1978).
16. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283**, 220-224 (1980).
17. Rogers, J. H. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1877-1879 (1980).
18. Ricciardi, R. P., Miller, J. S. & Roberts, B. E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4927-4931 (1979).
19. Maat, J. & Smith, A. J. H. *Nucleic Acids Res.* **5**, 4537-4545 (1978).
20. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1980).

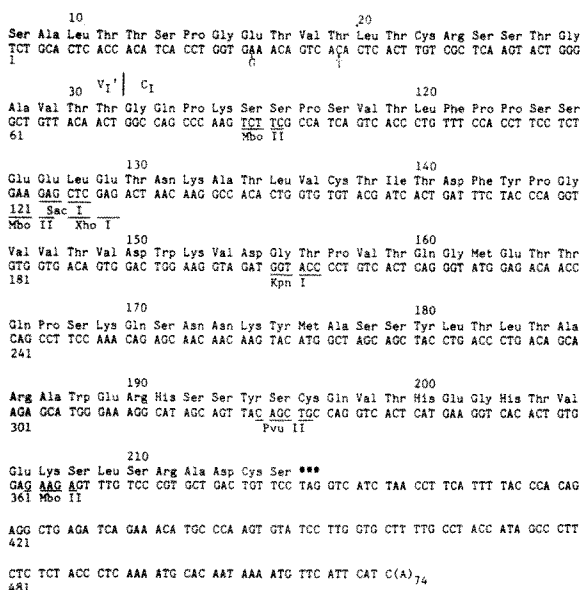


Fig. 3 Nucleotide sequence of p λ I-13. The sequence of the insert was determined by both the Sanger dideoxynucleotide extension procedures according to Maat and Smith¹⁹ and the Maxam-Gilbert methods²⁰. The sequence of both strands was determined by kinase treatment at the following sites: *Pst*I, *Xho*I, *Pvu*II and *Mbo*II (at position 207). The two residues below the sequence (G, T) are the only differences between the λ I and λ II variable regions in this clone. The numbers above the translation sequence are the amino acid positions in the mature protein and the numbers below the DNA sequence indicate the number of base pairs from the first nucleotide in the inserted λ I sequence as it is drawn here.

Recombination suppression of mouse *t*-haplotypes due to chromatin mismatching

Lee M. Silver* & Karen Artzt

Sloan-Kettering Institute for Cancer Research, New York, New York 10021, USA

We show here that recombination occurs at the normal rate in compound female mice containing two different complementing lethal haplotypes (t^{h17}/t^{w12}) where there is a long stretch of homologous *t*-mutant chromatin. Thus the recombination suppression of a complete *t*-haplotype cannot be due to an intrinsic factor(s) which suppresses along the length of its own chromosome but is due to 'mismatching' of wild-type and mutant chromatin. Naturally occurring *t*-haplotypes of mouse chromosome 17 have several interesting genetic properties. First, they are always transmitted from males in much higher proportions than mendelian expectation; presumably this accounts for the maintenance of lethal and semilethal *t*-haplotypes at polymorphic levels in populations of wild mice. *t*-Haplotypes also show recombination suppression. The conventional map distance between genetic markers *T* and *tf* is 7–12 cM, whereas in (*t*/+) heterozygotes for naturally occurring *t*-haplotypes, recombination is suppressed and *T* and *tf* seem to be separated by only 0.1–0.5 cM (ref. 1). The region of recombination suppression extends to and includes the major histocompatibility complex (H-2)². Thus *t* and H-2 effectively travel as a single unit—a 'super gene'³. Although recombination suppression is known to be accompanied by failure of chiasmata formation⁴, the mechanism underlying the suppression has remained an enigma. Lyon suggested a disorder of *t*-heterochromatin⁵ and more recently a change in 'intercalary' middle repetitive DNA⁶. She proposed that either *t*-chromatin is intrinsically incapable of participating in crossing-over, or chiasma formation is prevented because of mismatching and mispairing of normal and abnormal chromatin. We have measured recombination between two chromosomes which carried extensive overlapping segments of *t*-chromatin. We report here that in this configuration, recombination occurs at a normal rate, and thus we conclude that cross-over suppression is due to mismatching.

Rare recombination events do occur in *t*/+ heterozygotes and produce reciprocal cross-over products that are profoundly different. Whereas intact *t*-haplotypes contain a factor responsible for interaction with *T* to produce taillessness as well as a recessive lethal factor, those recombinants that retain the centromeric end of the parent *t*-haplotype continue to interact with *T* to produce taillessness, but are no longer lethal when homozygous, and those that acquire the distal end of the original *t*-chromosome retain the lethal factor but no longer show any interaction with *T*. Both cross-over products show limited recombination suppression in heterozygotes: *t*-viable haplotypes suppress it about half-way through the *T*-*tf* interval but not beyond⁷; *t*-lethal haplotypes seem to continue to be inextricably linked to H-2 although they permit free recombination between *T* and *tf*⁸.

Lyon investigated the problem of cross-over suppression by measuring recombination between recombinant chromosomes that contained, in one case, the proximal end and in the other the distal lethal factor, and which were thought to have a short overlapping region of *t*-chromatin. Using t^{h2} and t^{h17} (see Fig. 1a) she found very low recombination (1.5% compared with 18.1% expected for $t^{h17}/+$), and concluded that each *t*-haplotype suppressed recombination over its own length. She also

tested several haplotypes that are thought to be non-overlapping, and found that crossing-over was normal where the wild-type chromatins coincided (Fig. 1b, for example). These experiments did not rule out the possibility that cross-over suppression is due to mismatching of *t* and wild-type chromatin.

The characteristics of the two chromosomes we used (see Fig. 2) were as follows. The t^{h17} chromosome was derived by Mary Lyon⁸. It carries *T*; t^{h17} , an isolated lethal factor derived from t^6 (t^6 complementation group); $+^{tf}$ and translocation 190 (T190) inseparably linked to the lethal factor. *Trans* to wild type, it allows recombination in the region of homozygous wild-type chromatin between *T* and t^{h17} which we measure in males as 16.6% (193/1,161) between *T* and *tf*. This agrees well with Lyon's figure of 18.1% for males and females⁶. The $t^{w12}tf$ chromosome is a complete lethal haplotype containing t^7 (the mutant tail interaction factor), t^{w12} lethal factor (t^{w1} complementation group) and *tf*. The *tf* on this chromosome appeared spontaneously either as a point mutation or small deletion; recombination was not involved because the t^{w12} chromosome still retains its original H-2 type (Craig Hammerberg, unpublished data). We maintained the $t^{w12}tf$ chromosome in the balanced lethal cross ($Ttf/t^{w12}tf \times Ttf/t^{w12}tf$) homozygous for *tf* and thus the recombination rate is difficult to estimate accurately (see Table 1, line 2). However, recombination from the parental $t^{w12}+$ chromosome where the exchange of tufted can be monitored is 4.4×10^{-3} (24/5,576) (see Table 1, line 3). Taken together, the recombination rate between *T* and *tf* in the presence of $t^{w12}tf$ or $t^{w12}+$ is $<0.5\%$.

In choosing the two haplotypes t^{h17} and $t^{w12}tf$, we also took into consideration the fact that they share the same H-2 type, as this suggests that the two chromosomes may have a common descent from a single ancestral *t*-haplotype by mutation, and thus that their abnormal chromatins would be homologous.

Recombination measured in $t^{h17}/t^{w12}tf$ females was an extremely high 18.4% (Table 1, last cross). This is remarkably close to the value of 17–18% for t^{h17} *trans* to wild type, and high for the interval *T*-*tf*. Clearly, recombination is occurring freely between the t^{h17} and $t^{w12}tf$ chromosomes.

Of the 14 recombinant chromosomes obtained, 7 have so far been analysed. The important question was whether the two reciprocal types of recombinants, short-tailed tufted (*Ttf*) and normal-tailed non-tufted ($t^T + ^{tf}$ T190) carried one, or both, or

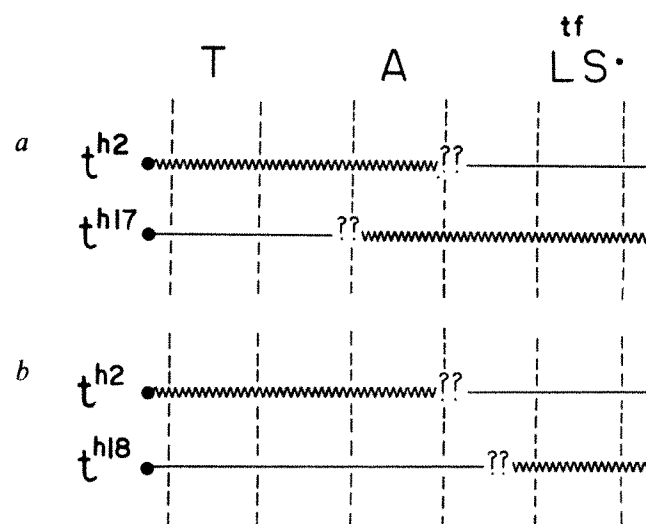


Fig. 1 Situations in which Mary Lyon measured recombination; diagrammed after Lyon *et al.*^{6,8}. *T*, the region of the Brachyury mutation (*T*), and the tail interacting factor (t^T) present in complete lethal *t*-haplotypes. *A*, as denoted by Lyon, the region of a ratio factor. *LS*, the region of lethality and sterility factor (*s*) is also the region where the marker tufted (*tf*) maps. *a*, Recombination between chromosomes with a short overlapping region of *t*-chromatin; *b*, recombination in haplotypes thought to be non-overlapping.

* Present address: Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

Table 1 Recombination data for T - tf interval in t^{h17} and t^{w12} mice

Cross	Normal tailed tufted	Normal tailed non-tufted	Short tailed tufted	Short tailed non-tufted	Tailless tufted	Tailless non-tufted	Recombination rate (%)
$\delta Tt^{h17}T190 +^{tf}/++ \times \delta +/++ \times tf/++ \times tf$	405	94*	99*	563	NE	NE	16.6†
$\delta Tt^{w12}tf \times \delta Tt^{w12}tf$	NE	16*	NE	NE	6,682	NE	0.24
$\delta Tt^{w12} +^{tf} \times \delta Tt^{w12} +^{tf}$	NE	20*	NE	NE	4.5‡	5,576	0.44
$\delta Tt^{h17}T190 +^{tf}/t^{w12}tf \delta +/++ \times tf/++ \times tf$	27	7*	7*	35	NE	NE	18.4

Mice were scored for tail length at birth and for tufted at 28 ± 2 days. The second and third group of crosses are balanced lethals where, in the absence of recombination, the only expected phenotype at birth is tailless; T/T and t^{w12}/t^{w12} embryos die before birth. All the normal tailed mice born were genetically analysed for their recombinant chromosome. Those that were fertile all proved to be $t^{w12}/t^{viable}tf$, that is carrying the proximal end of t^{w12} haplotype with the tail interaction factor but having lost the distal lethal component. The tailless tufted mice in the third cross represent the same class of recombinant but heterozygous with $T(Tt/t^{viable}tf)$. The reciprocal recombinant, $Tt^{lethal}+$, presumably goes undetected in this cross because it cannot meet a chromosome with which it is viable. NE, not expected.

* Recombinant phenotypes in each cross.

† The numbers in this cross include previously unpublished data of L. C. Dunn and D. Bennett, 1965–1973, and current data up to 1979.

‡ The 4.5 tailless tufted is a projected number based on one tailless tufted out of 1,242 tailless mice that were raised and scored for tufted.

Table 2 Analysis of new recombinant haplotypes

Haplotype	Cross	Complementation data			Recombination data				Rate (%)
		Normal tail	Short tail	Tailless	Normal tailed non-tufted	Normal tailed tufted*	Short tailed non-tufted*	Short tailed tf	
$Tt^{s1}tf$	$\frac{Tt^{s1}}{++} \times \frac{++}{Tt^0}$	83	33	12	7	0	1	8	6.3
	$\frac{Tt^{s1}}{++} \times \frac{++}{t^{w12}}$	142	23	2†					
$Tt^{s5}tf$	$\times t^{0+}$	28	36	11	21	3	5	75	7.7
	$\times t^{w12}$	21	28	0					
Tt^{s6}	$\times t^0$	12	26	16	37	1	2	54	6.1
	$\times t^{w12}$	11	29	0					
Tt^{s10}	$\times t^0$	12	6	3	41	4	6	84	7.4
	$\times t^{w12}$	27	0	0					
Tt^{s11}	$\times t^0$	15	35	14	31	5	4	54	9.6
	$\times t^{w12}$	95	15	2†					
$t^{s8}T190$	$\frac{T}{t^0} \times \frac{T}{t^0}$	0	NE	17			No data		
$t^{s8}T190$	$\frac{T}{t^{w12}} \times \frac{T}{t^{w12}}$	3	NE	2					
$t^{s9}+$	$\times t^0$		No data				No data		
	$\times t^{w12}$	5	NE	2					

The new recombinant haplotypes were arbitrarily designated $t^{s1,s2}$ and so on in order of their detection. For the complementation data, the numbers at birth contain pooled male and female data (except for $Tt^{s10} \times t^{w12}$) and are usually distorted reflecting the male ratio. Recombination data are obtained from crosses of $Tt^{s1}tf/++ \times +/++ \times tf/++$. Data are pooled reciprocal crosses except in the case of Tt^{s11} where the data are from females only. The numbers scored reflect preferential culling at birth against normal tails. NE, not expected.

* Recombinant phenotypes.

† This class in Tt^{s1} and Tt^{s11} are recombinants; three tested as recombinants, one never bred.

‡ Same as for t^{s1} crosses above.

neither of the lethal factors present on the parental chromosomes. The complementation data shown in Table 2 demonstrate that all five Tt^{s1} recombinants analysed carry the t^{w12} but not the t^{h17} lethal factor.

The data for the $t^{s1} +^{tf}$ T190 recombinants are still sparse, as the T190 translocation renders them semi-sterile, and difficult to analyse in complementation tests. Nevertheless, available data (Table 2) suggest that this class contains the t^{h17} lethal factor but not t^{w12} . At face value these data suggest that t^{w12} and t^{h17} act as alleles. However, three alternative and equally likely interpretations must be considered. First, t^{w12} and t^{h17} may be very closely linked, and only the analysis of many recombinants could define them as genetically separate. Second, all the recombination may be taking place in a restricted 'hot spot' to the left of both t^{w12} and t^{h17} . There is some evidence from exceptional recombination in mice heterozygous for t -chromatin that the second possibility is true. In eight exceptional recombinants from balanced lethal stocks containing various complete t -haplotypes and also an interstitial marker qk between T and tf ($T(3)qk(4)tf$), all eight cross-overs were between qk and the t

lethal factor (K.A., unpublished data). In the present study another marker ($Tcp-1$), known to be between qk and the t lethal factor⁹, was monitored; 10/10 recombinant chromosomes tested resulted from cross-overs between $Tcp-1$ and t^{lethal} . A third, obvious possibility is that the lethal factors are actually located distal to $tufted$, and as we screen for recombinants using $tufted$, they will always be exchanged.

Recombination data for the new haplotypes $Tt^{s1,s5,s6,s10,s11}$ (all of the $Tt^{w12}tf$ type) are also presented in Table 2. Clearly they now all allow recombination, just as the parental haplotype t^{h17} does, in the region of homozygous wild-type chromatin between the centromere and the t^{w12} lethal factor. All nine such recombinants tested were between T and t^{w12} . Recombination in these heterozygotes, however, is about 8–11%, considerably less than that from the Tt^{h17} T190 chromosome; this supports the suggestion of Lyon⁶ that the very high recombination in the translocation heterozygotes was produced by the translocation itself.

The transmission ratio data for the various recombinant haplotypes are given in Table 3. Clearly the Tt^{s1} chromosomes

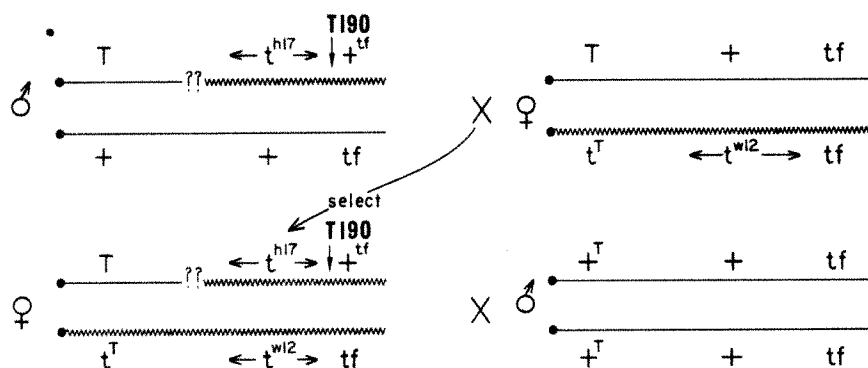


Fig. 2 Construction scheme of compound heterozygotes for t^{h17} and $t^{w12}tf$ and their expected offspring when mated to normal tail tufted. Short-tailed t^{h17} carriers were mated to tailless $t^{w12}tf$ mice. The tailless offspring, which are $Tt^{h17}/t^{w12}tf$, were selected. Males were tested for fertility. 4/4 were sterile, as expected, because they carried two complementing lethal factors. Females were mated to normal tailed tufted homozygotes to observe recombination between T and tf . Parental types: short tail, non-tufted, T190, and normal tail, tufted. Recombinant types: short tail, tufted, and normal tail, non-tufted, T190.

have a generally high, but very variable, ratio whereas the reciprocal type have very high and quite invariant ratios. For comparison note that the parent $t^{w12}tf$ chromosome is also transmitted in very high proportion (the last six males of this type we have tested with progeny samples over 25 averaged 98% t -bearing offspring), with individual male ratios of 100, 100, 100, 99, 96 and 87% respectively. The parent Tt^{h17} chromosome, on the other hand, in similar samples, gave an average ratio of 65%, and individual males were more variable, having ratios of 76, 73, 73, 72, 61 and 42%. Thus, recombination between the $t^{w12}tf$ and Tt^{h17} chromosomes has altered the transmission ratio characteristic of the respective lethal factors, implying that some element(s), such as the 'A' factor predicted by Lyon¹⁰ to lie between T and t^{lethal} , has been switched by recombination. This explanation does not completely account for the data, however, as the Tt recombinant ratios are both generally higher and more variable from male to

male than those typical of Tt^{h17} . Further work on interactions between A factors and specific t -lethal factors will be necessary.

The possibility remained that the translocation inextricably associated with t^{h17} was in some mysterious way responsible for overcoming the recombination suppression of the complete t -haplotype $t^{w12}tf$. Preliminary data rule out the presence of T190 as a factor in overcoming the suppression: one of the chromosomes derived from the above cross, $Tt^{h17}tf$, genetically analysed to be $Tt^{w12}tf$ and now devoid of T190, was set up *trans* with the complete haplotypes $t^{w5}+$ and $t^{w32}+$. Recombination scored so far in these compound females is 2/25 for t^{w5} and 6/33 for t^{w32} , demonstrating that the presence of the translocation is not required for recombination.

We set out to test the hypothesis that recombination in t -haplotype/wild-type heterozygotes is impaired because of mismatching between normal chromatin and abnormal t -chromatin. In fact, as all t -haplotypes are thought to be derived from a common ancestral haplotype¹¹ with divergences due to mutation in recombinationally isolated chromatin, whatever is wrong with the DNA in one t -haplotype should be matchingly wrong in complementing but mutant t -DNA. Thus, it is not unreasonable that variant t -DNA should recombine with variant t -DNA. Therefore crossing-over should occur in long regions of homozygous abnormal t -chromatin. Our data show unequivocally that recombination suppression is indeed abrogated in such combinations. We conclude that recombination suppression is not due to some intrinsic inability of t -chromatin to undergo crossing-over. An alternative hypothesis suggested above and consistent with the data is that mismatching between normal and variant chromatins is responsible. DNA cloning should show whether the non-coding sequences of t -DNA are rearranged or scrambled with respect to wild type.

All the recombination events we measured occurred in a restricted region proximal to the tf marker. Our experimental design, which used two different t -haplotypes with the same $H-2$ haplotype, precluded studying recombination distal to *tufted*. Nevertheless, it is important to know that recombination suppression associated with t -haplotypes can be circumvented, and the use of second-generation recombinants with different $H-2$ types will make possible dissection of the t -region identifying the number of genes involved, and mapping them with respect to one another.

K.A. thanks Drs Dorothea Bennett and Paulette McCormick for valuable suggestions. This work was supported by grants from the NIH.

Table 3 Transmission ratio of recombinant haplotypes

Haplotype	Male no.	Total informative offspring	% Of offspring carrying Tt or t -haplotype (ratio)
Tt^{s1}	1, 2	155	44
	3	127	28
Total		282	37
Tt^{s5}	1	63	64
	2, 3, 4	91	67
	5, 6	67	88
Total		221	72
Tt^{s6}	1	60	95
	2	119	89
	3	76	70
Total		255	84
Tt^{s10}	1, 2	110	51
	3	67	69
	4, 5, 6	88	69
	7	78	100
Total		343	70
Tt^{s11}	1	92	79
	2	69	35
Total		161	60
t^{s2}	1-2	9	100
t^{s3}	1-2	15	100
t^{s8}	1-7	146	88
t^{s9}	1-2	25	84
t^{s12}	1-6	83	84

Transmission ratio was measured in Tt haplotypes by mating $Tt/++$ males to $+/+$ females and scoring tails at birth. The ratio of Tt is calculated as short tail/total mice born. Occasionally the ratio of two or three brothers was measured collectively. The ratio of t -haplotypes was measured in two ways. First, $+/t$ males were mated to $T/+$ females, where the measure of the male ratio is tailless/short tailed plus tailless, as the normal tails are noninformative. Second, T/t males were mated to $+/+$ females and the ratio calculated as normal tail/total mice born. The data for the t -haplotypes are pooled within a haplotype because the numbers are scant because of semi-sterility due to the presence of T190.

Received 15 September; accepted 16 December 1980.

- Bennet, D., Dunn, L. C. & Artzt, K. *Genetics* **83**, 361-372 (1976).
- Hammerberg, C. & Klein, J. *Genet. Res.* **26**, 203-211 (1975).
- Snell, G. *Folia Biol., Prague* **14**, 335-358 (1968).
- Forejt, J. *Folia Biol., Prague* **18**, 161-170 (1972).
- Lyon, M. F. & Meredith, R. *Heredity* **19**, 301-339 (1964).
- Lyon, M. F., Evans, E. P., Jarvis, S. E. & Sayers, I. *Nature* **279**, 38-42 (1979).
- Bennett, D., Artzt, K., Cookingham, J. & Calo, C. *Genet. Res.* **33**, 269-277 (1979).
- Lyon, M. F. & Bechtol, K. B. *Genet. Res.* **30**, 63-76 (1977).
- Silver, L. M., White, M. & Artzt, K. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6077-6080 (1980).
- Lyon, M. F. & Mason, I. *Genet. Res.* **29**, 255-266 (1977).
- Dunn, L. C., Bennett, D. & Cookingham, J. *J. Mammal.* **54**, 822-830 (1973).

BOOK REVIEWS

Séveso: the intrigue and the infighting

Alastair Hay

THE message of *The Superpoison* is both powerful and moving. The third book on dioxin to have been published recently, it covers ground similar to a previous one by Thomas Whiteside (for review see *Nature* 283, 229; 1980). Margerison, Wallace and Hallenstein, however, are specifically interested in the effects of dioxin on the residents of Séveso. This small Italian town fifteen miles north of Milan was contaminated by the inadvertent discharge of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin (dioxin) when a reactor manufacturing 2, 4, 5-trichlorophenol overheated. Margerison *et al.* have written a well researched book on Séveso which deals with the subject at every level: human, scientific and political. Like Whiteside, they show that the Séveso residents were pawns caught in a process which rapidly lost its human element.

Confusion became a fact of life for the people in Séveso and the surrounding towns. Some leading scientists warned residents that dioxin was extremely dangerous; others, of equal repute, claimed there was no hazard. As the authors point out even the Catholic Church had to have its say. This was particularly true on the question of abortion brought into focus by the teratogenic properties of dioxin. When women, pregnant at the time of the reactor accident, were warned that exposure to dioxin might cause them to give birth to malformed children, the Vatican newspaper *Osservatore Romano* declared: "Is the mere possibility of the risk of malformation a sufficient reason to eliminate an unborn human being which undoubtedly has a greater chance of being born normal than abnormal?". Cardinal Colombo, Archbishop of Milan, joined the argument, according to Margerison and his co-authors, "by praising certain couples who had volunteered to adopt any malformed children born in the Séveso area, and appealing for others to come forward".

In the opinion of the authors it was the pregnant women from Séveso who had the most harrowing time. They claim that it was not only the church and politicians of a right-wing persuasion that made life difficult for the women, but that the Left were equally guilty in this regard, manipulating the situation at Séveso to advance their cause for abortion law reform in Italy. The campaign to reform the law and make abortion easier was

The Superpoison. By T. Margerison, M. Wallace and D. Hallenstein. Pp.236. (Macmillan, London: 1981.) £7.95.

eventually successful, but action came too late to help the women of Séveso. For many, the risk of having a malformed child was too great. The government claimed that in such cases abortion was legally available in Italy, the law making adequate provision. The reality was quite different. Most doctors refused to terminate pregnancies, subjecting the women to humiliating treatment. In the end only 34 pregnancies were terminated legally. At least twice that number are said to have been performed either abroad or illegally in Italy.

Controversy still rages in Italy over the effects of dioxin on the children born to mothers who decided against abortions. Three sources — the Italian Parliamentary Commission report on Séveso, the former head of the Health and Epidemiology Commission for Séveso, Professor Gaetano Fara, and Dr Guiseppi Reggiani, Director of Medical Research for Hoffmann-La Roche — all claim that there is no increase in the rate of malformation in the children born in Séveso. The authors of *The Superpoison* are not convinced. Acknowledging that the lack of statistics for malformation rates in Italy as a whole makes difficult the interpretation of the situation at Séveso, they nevertheless feel that there is "cause for serious concern".

Their concern is based on claims made by a local general practitioner working in Séveso. According to Alberto Colombi and the People's Scientific and Technical Committee — a largely left-wing association of doctors and scientists — there is gross under-reportage of the number of malformations. It is certainly true that the number of abnormalities is higher than official figures would suggest. The reason for their exclusion from published statistics, however, is explicable. Many of the malformations are minor. Their inclusion, although increasing the total number of cases, still does not alter the general pattern. Whichever set of figures is chosen and compared with the norm for various regions of Italy or with other countries in Western Europe, the answer is the same:

the rate for Séveso does not appear to be higher than anywhere else. However, research in progress may alter this conclusion.

The authors make much of the differences in the published figures for the number of cases of chloracne — the disfiguring skin disease caused by dioxin. In their view, "with so much disagreement between published figures, it is difficult to have confidence in the statistics" for the number of cases of chloracne observed. Again, the authors believe that the authorities are underestimating the seriousness of the situation and they cite various local doctors to support this charge. One such, the local Séveso health officer, Dr Francesco Uberti, is quoted as saying that he found 150 cases in children at one school alone. Uberti claims that the Milan regional authorities denied that his cases were genuine chloracne and "manipulated the facts". Diagnosis of chloracne can be difficult, as the authors admit, and this is at the basis of the disagreements about incidence. But this book was prepared some time ago, and the situation is now much clearer. A final figure for the total number of children who contracted chloracne has been agreed; it is 187.

Disagreements about the actions taken by the local authorities to protect the population at Séveso are well known. Margerison *et al.* insist that Givaudan and Hoffmann-La Roche, the principal companies involved in the dioxin problem, were aware of the dangers facing the residents of Séveso. They are unimpressed by the companies' arguments that it took time to construct a map of dioxin contamination preparatory to evacuating a section of the Séveso population. Indeed, they imply that the companies convinced the Italian authorities that the situation at Séveso was not dangerous. Two weeks after the dioxin discharge from the reactor, the authorities were indeed claiming that an evacuation was unnecessary. According to the authors, this complacency was shattered when Roche's Director of Medical Research insisted that an evacuation was absolutely essential. Many of these facts are not in dispute; however, the authors' thesis that the authorities were lulled into a sense of false security by Givaudan and Roche only to be caught out by Roche later on is not really proven.

All three authors are professional

journalists. Tom Margerison is a well known science journalist and this book has benefited from his skills. Technical details essential to the arguments are put clearly and simply. The operation of the trichlorophenol reactor at Séveso is well described. Accounts on the operation of the reactor differ, say the authors, depending on whether it is the workers themselves describing the operation or the management describing the instructions given to the workforce.

No satisfactory explanation is available yet to explain why the accident occurred. The authors find attractive a theory that there may have been sabotage. They say that one piece of evidence supports this. The reactor contents at Séveso are not alkaline as might have been expected. It is as though they were neutralized with acid, say the authors, giving credence to the sabotage theory. They do admit, however, that the pH change may be attributable to temperature differences, as indeed it is. Experiments designed to re-create the conditions in the Séveso reactor have since shown that at temperatures in excess of 260°C the pH of the contents changes from

alkaline to acid. The temperature of the reactor at Séveso is thought to have reached at least 350°C. Sabotage is ruled out in this case.

Political intrigue was rife at Séveso and this book describes it in ample detail. Scientific infighting also occurred and the authors take many irreverent swipes at the various experts involved with the dioxin problem. Interviews with many of the Séveso residents confirm that their treatment by the various authorities has been high handed. The authors succeed in showing that the residents at Séveso have had a raw deal. In their view the final immorality of the whole episode was that "Hoffmann-La Roche had the benefit of a profitable product while the innocent citizens of Séveso unknowingly took the risk". This may well be the verdict of history. But as more information accrues it is clear that the issue is not as clear cut as the authors make out. The companies certainly have much to answer for, but so too do the Italian authorities. □

Alastair Hay is a Lecturer in the Department of Chemical Pathology at the University of Leeds.

More from Harris

Alan E.H. Emery

The Principles of Human Biochemical Genetics. By Harry Harris. 3rd Edn. Pp.558. (Elsevier/North Holland Biomedical: 1980.) Hbk \$65.75, Dfl.135; pbk £26.75, Dfl.55.

THE first edition of this book, published in 1959, was an immediate success. The subject of human biochemical genetics was then in its infancy but there were sufficient pointers to indicate the exciting possibilities for the future. It has to be remembered that until Garrod introduced the idea of inherited disorders involving chemical processes (so-called inborn errors of metabolism) at the turn of the century, human genetics had been largely concerned with the inheritance of obvious structural abnormalities. In 1923 Garrod was only able to describe some half-dozen rare inborn errors of metabolism, but now Professor Harris details some 120 such disorders and a further 73 enzyme and protein polymorphisms.

This new edition is divided into nine chapters with an appendix listing details of specific enzyme deficiency disorders, and a similar appendix of enzyme and protein polymorphisms. The introductory chapter on gene mutations and single amino acid substitutions is, like the rest of the book, a model of clarity and could well form the basis of any graduate course in genetics. Genes can no longer be considered merely discrete sections of DNA contiguous with each other, but are frequently separated by intervening sequences where the relevant messenger RNA is not translated into protein. The function of these intervening sequences is as yet not clear, but in the future the unravelling of their role in gene action could well herald yet another new departure in biochemical genetics.

Two chapters are devoted to the principles of gene action which are largely, though not exclusively, illustrated from findings in the haemoglobinopathies. A discussion of gene localization (mapping) is also included. Until recently gene localization was dependent on the study of pedigree and somatic cell hybrids, but is now being greatly helped by the use of restriction endonucleases. In the near future this, too, is going to generate an entirely new perspective of gene structure and localization.

Subsequent chapters also deal with quantitative and qualitative variations in enzymes, inborn errors of metabolism and blood group substances. However many enzyme and protein variations (and the more recently discovered variations at the DNA level of restriction endonuclease cleavage sites) appear to be unrelated to any specific disorder, yet often their incidence is far greater than could be accounted for by mutation alone. Harris suggests that the term polymorphism might be reserved for those variations in which

Plasma physics for robust readers

A.D.R. Phelps

Plasma Physics for Nuclear Fusion. By Kenro Miyamoto. Pp.610. (MIT Press: 1980.) £35, \$50. *Fundamentals of Plasma Physics.* By V.E. Golant, A.P. Zhilinsky and I.E. Sakharov. Pp.405. (Wiley-Interscience: 1980.) £22.10, \$51.90.

BOTH of these books treat several aspects of plasma physics in detail, and yet neither provides a complete coverage of the subject — Golant *et al.* intentionally omit the important subject of waves in plasmas and Miyamoto, as implied by his title, concentrates on those topics in plasma physics which relate to nuclear fusion. Both are appropriate for postgraduate courses, and should appeal particularly to the mature research scientist.

The text by Miyamoto uses rationalized mks units and is a translation of his 1976 Japanese publication. Some opportunity appears to have been taken to update a few of the references, but some of the most recent results of nuclear fusion research have not been included. Nevertheless, much of the content has not hitherto existed in such digestible textbook form.

Professor Miyamoto divides his 16 chapters into four sections: fundamentals, which includes a review of fusion criteria, magnetic field configurations, single particle motion, Coulomb collisions and kinetic equations; MHD, covering equilibrium, diffusion and confinement, and instabilities; kinetic description of waves and instabilities, which treats waves

in cold and hot plasmas and velocity space instabilities; and finally, heating, diagnostics and confinement.

Fundamentals of Plasma Physics is also a translation (from the original 1977 Russian edition) and, like Miyamoto's book, more readable than might be expected. In ten chapters the authors cover collisions, kinetic equations, equilibrium plasmas, distribution functions, transport processes and magnetic field influences on particle motion, transport processes and plasma confinement. Cgs units are used. The notation is somewhat inconsistent in places, for example rot and curl appear on the same page and the Laplacian operator used in the text does not correspond to that given in the symbols list. Yet these are minor points which should not worry the relatively robust reader for whom both books are best suited.

Whereas Miyamoto's book essentially involves the high temperature collisionless type of plasma, Golant *et al.* address themselves more to gas discharges and ionization phenomena. Although there is some overlap in subject matter between the two books, the difference in emphasis makes them complementary. Thus one might expect the libraries of those research laboratories and universities engaged in plasma physics research to invest in both of them. □

A.D.R. Phelps is a Lecturer in Natural Philosophy at the University of Strathclyde.

the commonest identifiable allele has a frequency no greater than 0.99 and therefore the proportion of heterozygotes is at least two per cent of the population. It is difficult to account for the high prevalence of such polymorphisms and the author presents a very readable account of the problem. It might, of course, turn out that most of these variations are no more than "random noise", but it is difficult to accept that this could be the entire explanation.

Altogether this is a fine book which can be unreservedly recommended as a graduate student text as well as a work of

reference. Of course, because of the enormous growth of the subject, this new edition might now appeal more to the specialist. Perhaps for a more general readership Professor Harris could be persuaded to produce a basic text similar to his *Introduction to Human Biochemical Genetics* which was published some 22 years ago by Cambridge University Press and which was the progenitor of the current text. □

Alan E.H. Emery is Professor of Human Genetics in the Medical School of Edinburgh University.

Where psychology meets biology

M. J. Morgan

Essay on Mind. By D. O. Hebb. Pp.159. (Lawrence Erlbaum Associates: 1980.) \$16.50.

IN THIS "essay", Professor Hebb expresses in clear and forceful language his conviction that "... mind and thought, consciousness and creativity and free will, are all biologically evident phenomena and seen most clearly in the light of evolutionary ideas". He summarizes the historical development of his own "cell assembly" theory, familiar from his classic book *The Organization of Behaviour* (Wiley, 1949) and to a certain extent attempts to relate it to modern developments in neurophysiology and psychology. There is little here that is very new, but Hebb was one of the first to argue that psychology is largely a biological subject, at a time when it was neither profitable nor popular. The gritty style of this book makes few concessions to the subtleties of professional philosophy, and somehow manages to convey the general impression that objections would be unwise and ill-informed, if not actually silly. Fortunately, no one need feel too offended by this brusque treatment, because the counter-arguments are put into the mouths of anonymous "writers", rather than being attributed to specific people.

On the heredity-environment issue Hebb neatly clears up some common confusions, and expresses the view that extreme environmentalism can have morally dangerous effects. It is pleasing to see the boot on the other foot for a change: usually it is environmentalists who assume the mantle of moral superiority. Returning to a theme familiar from his *Textbook* (W. B. Saunders, 1972) Hebb warns the young student that if he has an inherited predisposition to alcoholism, he had better recognize it quickly and take appropriate avoiding action. This he is not likely to do "... if he takes it for granted that all men are born equal ... and seeing others drink freely without becoming drunks,

takes it for granted that he can do the same". This is probably good advice, certainly no worse than that given by people who think that it is wicked to believe in genes.

The "cell-assembly" model can now be seen as one of the first "feature analysis" models of perception. Its essence is that cells or groups of cells coding for distinctive features such as lines are grouped together into superordinate assemblies by associative learning. These assemblies can in turn be grouped into assemblies representing concepts, and so on, presumably up to "grandmother assemblies", which would be activated when we were thinking of a particular person. Hebb notes that the physiological work of Hubel and Wiesel has given considerable support to this idea since he first produced it, but he quotes their hierarchical model as if it were established fact, without mentioning important criticisms. He also attempts to extend the theory in a way that many readers will probably find frankly metaphorical: we read, for example, that "Emotion is a disruption of cell assemblies".

The experiments that Hebb pioneered at McGill on effects of early environment in animals are given prominence and still raise issues of great interest. But as in the case of the discussions of visual perception, it is disappointing not to be presented with any more recent data, particularly those concerning effects of early environment on synaptic connections.

Possibly the most interesting aspect of this book is the historical insight it gives into the process by which Hebb arrived at his theory at a time when radical behaviourism made any physiological theorizing disreputable. That many of these early battles will now seem a bit outdated to the younger generation of psychologists is a measure of Hebb's success in helping to win them. □

M. J. Morgan is Professor of Psychology at the University of Durham.

ISSN-0275-4207

THE BIOTECHNOLOGY NEWS SOURCE™



BioEngineering News -
the *only* biweekly
newsletter covering the
genetic-engineering
recombinant-DNA,
monoclonal-antibody,
and cell-culture
industries.

24 issues — \$236.00 (U.S.)
(\$195.00 in U.S. & Canada)

Subscriptions sent airmail.

Write: **BioEngineering News**, Suite
304, 109 Minna St., San Francisco,
CA 94105. Tel: (415) 332-0508.
Telex: 171463-HQLARK.

Special Reports Available: (1) Genetic Engineering Industry, (2) Hybridoma & Cell Fusion, (3) Automated DNA Synthesizers. Price: \$29.95 (U.S. Funds).

NOTE - For prepaid subscriptions select any two (2) reports as a bonus.

Circle No.16 on Reader Enquiry Card.

The ANS Journals

Your number one source for current information on the nuclear industry.

- NUCLEAR NEWS (NN)
 - NUCLEAR TECHNOLOGY (NT)
 - NUCLEAR REPORT (NR)
 - NUCLEAR STANDARDS NEWS (NSN)
 - ANSPI (ANSPI)
 - NUCLEAR SCIENCE & ENGINEERING (NS&E)
 - NUCLEAR TECHNOLOGY/FUSION (NT/F)
 - TRANSACTIONS (TR)
 - REMOTE SYSTEMS TECHNOLOGY (RSTD)
- (Write for our complete publications catalog.)

Send To:

AMERICAN NUCLEAR SOCIETY
555 N. Kensington Avenue
La Grange Park, IL 60525 USA

Please send me information on...

- ☐ NN ☐ NS&E ☐ NT ☐ NT/F ☐ NR
☐ NSN ☐ TR ☐ RSTD ☐ ANSPI
☐ National Meetings ☐ Topical Meetings
☐ Individual Membership ☐ Corporation Membership
☐ Send me your free catalog

Circle No.11 on Reader Enquiry Card.

NAME _____

TITLE _____

ORGANIZATION _____

ADDRESS _____

CITY _____

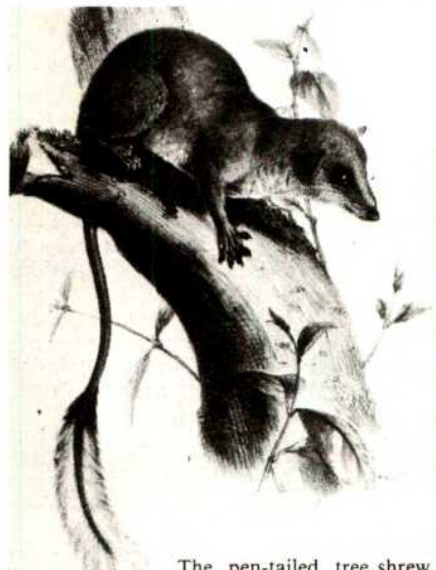
STATE _____ ZIP _____

Clearing the decks

Adrian Friday

Comparative Biology and Evolutionary Relationships of Tree Shrews. Edited by W.P. Luckett. Pp.314. (Plenum: 1980.) \$39.50, £24.89.

TREE shrews are placental mammals of Southeast Asia. Although not rodents, they resemble squirrels and the Malay word "tupai" is used for both. Of all placental mammals, tree shrews are probably the most enigmatic and differences of opinion concerning their evolutionary relationships



The pen-tailed tree shrew, *Ptilocercus lowii*, from the plate in the original description by J.E. Gray.

have caused some heated exchanges among mammalogists. Originally included in that ragbag of an order, Insectivora (within which they had been particularly allied with the elephant shrews of Africa), the tree shrews came to be widely regarded as the least-specialized grade of primate largely as a result of the work of Alberta Carlsson and W. E. Le Gros Clark. In this volume there are contributions on the tree shrew skeleton, dentition and cranial circulation, their meagre fossil record, their nervous and reproductive systems, and the molecular evidence for their relationships. Behavioural studies on tree shrews are not represented: distinguished work in this area has already been published, but Luckett implies that it has little to contribute to solving the problem of tree shrew relationships.

Contributors to the volume were asked to consider two questions in their accounts: first, do tree shrews and primates share unique derived characters suggesting an immediate common ancestry; second, are there reasons for supposing that tree shrews, primates, bats and colugos ("flying lemurs") together form a discrete evolutionary group (Archonta) amongst mammals?

It was part of the message from the work

of Van Valen and of Martin in the late 1960s that progress in assessing the evolutionary affinities of the tree shrews and other problematic groups, collected together in the past as Insectivora, would be made only if comparisons were made between these animals and all other mammalian orders. However, if Luckett had insisted on such application from all his authors, several of them would have been consigned to a lifetime's work.

The fascination of tree shrews reaches beyond biologists interested in Insectivora (in the widest sense) to include those concerned with the principles of phylogenetic reconstruction. In his introductory chapter, Luckett succinctly summarizes the phylogenetic methods described by Hennig, and uses them to dispose of the early arguments for immediate common ancestry shared by tree shrews and elephant shrews. Particular emphasis is given to the reconstruction of the primitive state for various characters in the hypothetical eutherian common

ancestor. In reviewing the evidence of all the contributors, Luckett reaffirms Butler's and McDowell's concept of the isolation of both the tree shrews and the elephant shrews, leaving other Insectivora as an assemblage of animals which may have some claim to monophyly.

It is hardly surprising that there are few firm conclusions concerning the relationships of the tree shrews, although Szalay and Drawhorn reassert the validity of Archonta, a tactic supported in varying degrees by the two chapters on molecular evidence. Inevitably, much recent work on tree shrews has been inconclusive, even destructive, but this might be viewed as a clearing of the decks. In any case, a volume that includes several chapters of thoroughly admirable compilation and evaluation of morphological data is very welcome for that alone. □

Adrian Friday is Assistant Curator of Vertebrates at the University Museum of Zoology, Cambridge, UK.

Pedagogical overview of quantum fields

Paul Davies

Quantum Field Theory. By Claude Itzykson and Jean-Bernard Zuber. Pp.700. (McGraw-Hill: 1980.) £26.95, \$55.50.

A NEW generation of textbooks on quantum field theory is appearing, and this monumental work of 700 pages promises to become the standard book to replace Schweber's *An Introduction to Relativistic Quantum Field Theory* (Harper & Row, 1961), or Bjorken and Drell's *Relativistic Quantum Mechanics* (McGraw Hill, 1964) and *Relativistic Quantum Fields* (McGraw Hill, 1965), both nearly 20 years old. The subject has moved a long way in that time, and a comprehensive, pedagogical book covering both the traditional material and modern developments is long overdue.

The authors have been lavish and ambitious in their approach. Starting with classical field theory, they build upon an extensive discussion of the Dirac equation towards a thorough treatment of conventional quantum electrodynamics. On the way, all the mathematical formalism needed for the more familiar applications is presented. Also included are some topics that are generally hard to find in textbooks: the Klein paradox, the Heisenberg-Euler effective action, the Fock-Schwinger proper time method, finite temperature states, massive vector fields and the Casimir vacuum force effect. Especially welcome are the treatments of symmetries, the PCT and spin-statistics theorems, and a section on Grassmann algebra.

All the routine machinery of interacting

relativistic quantum fields then emerge via the S-matrix and Feynman rules, with extensive applications to quantum electrodynamic processes on the way. Then, on through dispersion relations and analyticity to an extensive section on radiative corrections and renormalization theory. Once again, one welcomes the inclusion of modern topics like Wick rotation, asymptotic behaviour and dimensional regularization.

But all this comprises only the first half of the book! The really attractive feature is the comprehensive treatment of recent functional methods, gauge theories and asymptotics which occupies the last 300 pages. An extensive discussion of non-Abelian gauge fields and spontaneous symmetry breaking opens the way to topics such as the Weinberg-Salam theory, anomalies, Fadeev-Popov ghosts and scaling behaviour in scattering theory.

This valuable survey is written with meticulous care and thoroughness. Although intended primarily for graduate students, there is such a wealth of material that undergraduates will find it useful for reference. Those learning relativistic quantum mechanics will also benefit from the chapter on the Dirac equation, hole theory and hydrogen-like atoms. Certainly, any student of quantum field theory will find everything necessary for a thorough grounding in this increasingly important subject. □

Paul Davies is Professor of Theoretical Physics at the University of Newcastle upon Tyne.

nature

12 March 1981

How true is the theory of evolution?

So is Darwin's theory of evolution a fact, a pack of lies or something in between? This is the question prompted by the two letters with which the British Museum (Natural History) has broken its long silence over the rights and wrongs of cladistics and other matters (see page 82). That this issue has burst into the open in the same week as the trial in California of the suit by the Creation-Science Research Center of San Diego against the Californian State Board of Education is both ironical and apt. In California, the creationists are complaining that the theory of evolution is being taught as if it were a fact, and that the constitutional rights of Californian schoolchildren are thereby being infringed. The correspondents from the Natural History Museum complain that a leading article in *Nature* two weeks ago implied that the theory of evolution is a fact. Both sets of complainants deserve to be taken seriously, the people from the museum more seriously than the creationists. Both are mistaken.

The question they have separately raised is, nevertheless, interesting and important. The theory of evolution is now more than a century old, and has been much transformed and modified during that long period. The objectives of the theory are what they always were, however — to account for the diversity of species as well as for their similarities and differences, and also to account for the evidence of the fossil record, which suggests changes in the patterns of living things in the course of time. It is acknowledged that this apparently innocuous statement begs one question which some consider to be important. The word *species* implies that species exist. Some out-and-out cladists prefer to avoid this assumption. Many of the differences between apparently mutually infertile groups of plants or animals may, the cladists say, be differences of degree and not of kind. Although this issue may be important in its own right, it is not strictly relevant to the status of evolution as a theory. Cladistics, a technique of classification, is in no sense inconsistent with the theory of evolution. Indeed, on the face of things, cladistic systems of classification should be more easily accommodated with the theory of evolution than those based on species. Cladism, therefore, is irrelevant to this journal's quarrel with the Natural History Museum, at least as now defined.

Another begged question lies in the word *evolution*. It is widely agreed that the fossil record implies substantial changes from one time to another in the pattern of extant living things, but it is not self-evident that the relationship between one pattern and a later pattern is properly described as evolutionary. That term implies that the pattern of living things at any time is determined by the pattern of living things at any earlier time and by environmental influences in the interval. Technically, evolution is a stochastic process. Because the fossil record can be (and frequently has been) interpreted in different ways, it is important openly to acknowledge that the essence of Darwinism includes two propositions — first, that successive stages in the fossil record are indeed related to each other in an evolutionary fashion and, second, that the mechanism of the transition from one to another consists of natural selection. So much even the creationists accept. It is not the content but the status of Darwinism that is now disputed.

But is not the theory of evolution simply a theory, a kind of glorified hypothesis that may be knocked down at any time, either because it is falsified or because a better theory comes along? This is what the round robin from the Natural History Museum asks. (The creationists, of course, go further.) The issue is complicated, and much confused by over-hasty readings of what Sir Karl

Popper has written on the significance of scientific theories as such. Theories are, of course, putative explanations of phenomena, or facts. Popper's great achievement, first of all stated in "The Logic of Scientific Discovery", was to point out that theories are not constructed by some non-existent logical process called "induction" but, rather, by the exercise of creative imagination. In the same important work, first published in 1934, Popper drew a sharp distinction between two kinds of theories — those which are capable of being falsified, for example by comparison with experiment, and those which cannot be thus falsified. Theories of the first kind, Popper says, belong properly to empirical science. Theories of the second kind, those which cannot be falsified, he called metaphysical. One obvious example of such a theory is that the world was made by God. Another is Darwinism.

Within the terms of the definitions he has helpfully constructed, Popper is technically correct. The course of supposed past evolution cannot be rerun. Nor is it possible to make predictions about the pattern of life-forms on the surface of the Earth that will be verifiable within the lifetime of people now alive. (Predictions that there may be fewer whales in the southern oceans some years from now are not tests of the theory of evolution but of the resolution of the International Whaling Commission.) Nevertheless, the theory of evolution is not entirely without empirical support. Dzhzhansky's classic work on the course of speciation among birds in Central America is a demonstration that, among one group of animals, the course of change is evolutionary, and that the mechanism is natural selection. Nevertheless, such demonstrations do not constitute a proof of Darwinism. They do not, for example, exclude Lamarckian processes or episodes of supernatural creation in past times. Darwinism is not the only theory afflicted by such problems. Cosmological theories are similarly incapable of falsification and are thus, in Popper's language, metaphysical. But metaphysical theories are not necessarily bad theories. Plainly it would be absurd if Popper's work were to be held to invalidate theories of evolution in general.

So how, in these circumstances, is it possible to tell the difference between good and bad theories? The first requirement of any theory, good or bad, is that it should be consistent with such phenomena as there are, and logically self-consistent. Darwinism is consistent with the data to which Darwin had access more than a century ago. One of the remarkable features of the theory is that it remains consistent with the vastly greater body of data now available. Yet Darwinism might have been falsified by any one of the host of discoveries of the past century. Although there is a certain randomness in the way in which palaeontological findings come to light, so that it is not possible to say that this current of discovery amounts to a stringent test of Darwinism, surely it is too wooden to insist on a strict interpretation of Popper's criterion for telling when a theory belongs to "empirical science" and when it is merely metaphysical. There is nothing in the letters from the Natural History Museum to suggest that the multitude of authors do not agree with that modest assertion.

On the grounds of internal consistency, Darwinism has also triumphed quite remarkably. In Darwin's time, the phylogenetic relationships between different living forms were inferred from their morphology. At the outset, it was entirely unforeseen that there would some day be more objective ways of constructing these relationships. Now, however, molecular biology has provided an independent (and potentially extremely powerful)

method of telling the relationships between species and groups of species, many of them only distantly related to each other. The result is a striking confirmation of the general character of the relationships suggested by Darwin and his contemporaries. To be sure, it is not possible to tell from a comparison of the structures of proteins found in primates whether people are more closely related to gorillas or chimpanzees, but that only goes to fill out Darwin's notion of divergent evolution from some common stock. At the same time, however, the quite remarkable constancy of some materials, the histone proteins for example, is vividly suggestive of the common origin of all living things, and of their persistence over time. Creationists would, of course, insist that none of this constitutes a proof of Darwinism in the strict sense. (If the histones are that special, God may have had no choice but to use them over and over again.) But the way in which the theory of evolution has been able to survive such a long succession of discoveries bearing on the mechanism of inheritance — the rediscovery of Mendelism, the discovery of chromosomes, the recognition of what genes are and the recognition that genes are usually pieces of double-stranded DNA — is striking evidence of its overwhelming consistency. No theory of such a grand scope in the physical sciences has done as well in the past century.

Logically, however, none of this amounts to proof. Popper, the creationists, the Natural History Museum and everybody else agree. It is also fair to acknowledge that there are matters on which a better agreement between theory and sketchy data is desirable. Much has been made of the episodic nature of the fossil record — the apparently long persistence of some species, the apparently sudden appearance or disappearance of others. If, however, the internal consistency of Darwinism cannot be counted as a proof, the supposed "gaps" in the fossil record, which may be a consequence of sampling bias of the kind that besets cosmology, cannot be counted as a disproof. It is also true that too little is known of the molecular structure of chromosomes (chromatin) for the mechanism of speciation to be fully understood. The time scale of supposed evolution is another supposed difficulty. How, the creationists ask, can the acme of evolution so far — the evolution of hominids — have been accomplished in a mere million years or so when it is supposed to have taken a thousand times as long to produce the first multicellular organisms? Honest Darwinists will reply that there are at present very few means of making quantitative calculations of this kind but they may also point to the way in which molecular biology is showing how a few genes may determine a surprising range of functions. In other words, there is no particular reason to suspect that outstanding problems of agreement between qualitative expectation and inadequate data will not in the end be accommodated. Creationists would naturally disagree. The Natural History Museum would probably have an open mind.

But what of the ways in which Darwinism may have to be amended? In the past few months, for example, the experiments of Gorgczynski and Steele (see *Nature* 19 February) have fanned excitement among the sceptics because they suggest that Lamarckian inheritance (acquired characteristics and all that) may be possible, in the laboratory if not in real life. So, if those experiments are confirmed, would Darwinism be dead? On strict logic, giving the two components of Darwinism (evolution and natural selection) equal weight, the answer would be yes. From a more practical point of view, the answer would take the form "it depends". If there turned out to be a well-defined set of circumstances in which Lamarckian inheritance occurred, and if these circumstances were in some sense exceptional, then the modification of Darwinism that would be required would be comparable with Einstein's modification (in special relativity) of Newtonian mechanics. "Darwinism is dead, long live Darwinism" would be the shout. And it is inconceivable that the circumstances could be other than exceptional — the regularity of the phylogenetic relationships revealed by molecular biology may not be a proof of Darwinism but they give the lie to Lamarckian inheritance as the engine of evolution. But even these considerations are for the time being irrelevant; the experiments concerned, which are capable of other than Lamarckian

interpretation, have not been confirmed. Most probably, the signatories of the letter from the museum would agree. Creationists would say that the argument is irrelevant.

Why, then, are the museum and *Nature*, both moderate institutions, at loggerheads? The leading article complained of did not ask that the theory of evolution "should be presented as fact" but protested at the use of the phrase "If the theory of evolution is true . . ." in a document intended for public circulation. The same article asked scornfully whether ". . . the theory of evolution is still an open question among serious biologists?" Two separate points arise, the first a matter of presentation. Knowing that large sections of the general public are sceptical of Darwinism, can it be sensible to use the phrase "If the theory of evolution is true . . ." when the museum's judgement is that there is "overwhelming circumstantial evidence" for Darwinism? To do so is to avoid challenging the sceptics when this is not merely possible but necessary. While some who doubt Darwinism do so on respectable grounds, others claim that the course of events may be determined by literally supernatural influences. Theories of that type are not even metaphysical — they are simply unscientific.

Nobody asks that the museum should declare a faith in Darwinism that would be unwarrantable but merely that it should choose its words more carefully. As the museum's brochure now reads, there is nothing to help the innocent reader to distinguish between alternatives to Darwinism such as the argument due to Hoyle and Wickramasinghe that the Earth has been repeatedly repopulated with living things (which is in principle a scientific theory, but for the time being metaphysical in Popper's sense) and entirely unscientific theories such as those of the creationists. It is mystifying that the museum should not have discovered this danger for itself. That is why the ambiguity of its public brochure is so shocking.

The second issue to which the museum needs to pay attention concerns its place within the professional community. The museum is a distinguished centre of scholarship. Given its interests, the museum's scholars are bound from time to time to be caught up in controversy. Its flirtation with cladism, although entirely proper and permissible, shows what passions can inflame those who look to the museum as a partner in research. Such troubles are unavoidable. On the evidence of the recent past, and to its credit, the museum has not been afraid to push its collective view that cladism is a valuable technique for classifying organisms. So why should it be so diffident in its approach to the grand theoretical issues? That Darwinism may ultimately be falsified is beyond dispute, but — if the round robin represents a collective view — the museum's ostentatious agnosticism on the subject is potentially dispiriting for the rest of us. If, in the museum's view, Darwinism is "seriously an open question" in the sense that some of the doubts about Darwinism raised in recent years are taken seriously, then it should take the lead in throwing light on these disputed issues. Who else? Yet there is no sign that the museum's research programme differs radically from what it has been in recent years.

The trouble with agnosticism is that, however well justified logically, it can be carried too far. If young scientists are brought up to believe that by their essence all theories are either falsifiable or metaphysical, is there not a serious risk that regard for the scientific enterprise as a means of understanding how the world is made will be undermined? Their elders are similarly seduced into the belief that, because nothing can be true, grubbing for data is all that matters. Whatever the philosophical position, there is no way of denying that people's emotions are engaged by the strengths and weaknesses of theories. To acknowledge that is not to acknowledge human frailty but simply (following Popper) that the construction of theories is a creative process, constrained only by conventions intended to suppress both speculation and prejudice. In all this, however, one prejudice is allowable, even necessary — the preconception that theories can be constructed to account for all observable phenomena. That is what the creationists cannot accept. It would have cleared the air if the signatories from the museum had said where they stood.

Mixed welcome for genes on Wall Street

Genentech's interferon from yeast plasmid

Washington

Anybody interested in following the progress of interferon research may soon find it easier to refer to the pages of the *Wall Street Journal* than to look in the scientific literature. Last week, shares in the West Coast genetic engineering company Genentech Inc. rose \$7, from \$36 to \$43, following the company's announcement that collaboration between Genentech and scientists at the University of Washington in Seattle had yielded interferon from genetically altered yeast cells. A few days earlier, shares in another publicly quoted company, Flow General, fell from 32 to 29½ cents with the news that the company faced delays in delivering the conventionally produced fibroblast interferon which it has agreed to supply to the National Cancer Institute for clinical trials.

The Genentech announcement was made during the First Annual Congress of Recombinant DNA Research, a privately sponsored meeting held in San Francisco. It represents the first public announcement of the successful use of yeast to produce mammalian proteins, although several other universities and companies are also working in the field and have yet to make their results known.

The techniques were developed by scientists in the University of Washington's genetics department, which has been studying the genetics of yeast since 1950. Dr Benjamin Hall, chairman of the department and one of the leaders of the research team, feels that the use of yeast could improve methods of producing interferon by up to a factor of ten.

Genentech has subsequently announced that it expects to be marketing interferon "before 1985". The company's vice-president, Mr Frederick Middleton, told a meeting of security analysts in San Francisco that the company's new human insulin, being developed with Eli Lilly and Co., would probably be on the market by the end of 1983. This would be followed within a year by human growth hormone, a bacterially produced version of which is now being tested on 20 children at the Great Ormond Street Hospital for Sick Children in London (see *Nature* 5 March, p.6).

Despite the long-term prospects, Mr Robert Swanson, the company's president, warned the analysts not to expect impressive performance figures from the company before it begins to market products. Citing the need for heavy investment in research and development,

as well as clinical trials, Mr Swanson said that in the next 12 to 24 months "it mightn't be possible to show a profit on a quarter-to-quarter basis".

In contrast to its bullish reaction to Genentech's news about interferon production, Wall Street responded less kindly to the announcement from Flow General that the first delivery of human fibroblast interferon to the National Cancer Institute would be delayed because of contamination problems with the cells.

The company's biomedical subsidiary, Flow Laboratories, has signed a contract with the institute to provide 50,000 million units of fibroblast interferon — enough for 50,000 doses — using a technique invented and patented by the Massachusetts Institute of Technology in which the interferon-producing cells are grown on the surface of small charged spheres. According to a report issued by the company, the viability of scaled-up production using the so-called "super-beads" was demonstrated in December. Since then, however, contamination of the cells has resulted in a "setback of an indeterminate period" which will delay the deliveries to the institute, due this month, by "some weeks".

The institute is already sponsoring

clinical trials at three centres using leukocyte interferon supplied by Meloy Laboratories of Springfield, Virginia, and at two centres with lymphoblast interferon produced by the Wellcome Foundation in London. It is also about to invite bids for the supply of gamma "immune" interferon for possible clinical trials, which has attractive properties distinct from the other more easily produced interferons. Flow General has reported that it is evaluating a process developed at New York University for the production of gamma interferon using conventional cell systems, to see if it would be able to produce sufficient quantities for further evaluation.

There is also evidence that the market's previous enthusiasm for genetic engineering stocks is waning. When the Berkeley-based Cetus Corporation went public last Friday, its shares opened at \$23, fell to a low of \$22½, and ended the day at \$23¼. Although this represented a record \$119 million initial public stock offering for a new company on Wall Street, Cetus's debut was much less spectacular than that of Genentech, whose shares soared from \$35 to \$80 on the first day of trading, but have since dropped back to near their original price.

David Dickson

Benefits and snags of yeast plasmids

Until Genentech and the University of Washington, Seattle, announced their success with interferon, there had been little news, most of it gloomy, about the chances of expressing mammalian genes in yeast rather than bacteria. And this despite the establishment of at least one company, Collaborative Research Inc., specifically to exploit yeast. A consultant for Collaborative Research said last week that part of the problem was simply that not enough was known about the basic molecular biology of yeast to make easy its subversion for unnatural purposes.

One important setback for the advocates of yeast occurred early on, when it was discovered that yeast could not cope with split genes. Most mammalian genes have their coding sequences interrupted by non-coding, or intervening, sequences. Bacterial genes are not split in that way and, not surprisingly, bacteria do not possess the mechanisms to decode the information encoded in split genes. The hope was that yeasts, more highly evolved than bacteria, might possess split genes and therefore be able to cope with mammalian split genes. But this has not proved to be the case. Indeed, the overriding reason for the successful expression of the human interferon gene in yeast is because it is that human rarity

— an unsplit gene.

Nevertheless, advocates of yeast still have reasons to be cheerful:

- Even split genes produce unsplit messenger RNA (the intermediary between the chromosome and machinery of protein synthesis in the cell) from which unsplit DNA, suitable for inserting into yeast, can be made with increasing ease.

- Inasmuch as there is some redundancy in the genetic code, yeast tends to share with mammals certain preferences of codon usage that are not favoured by bacteria; consequently yeast may find it easier than bacteria to decode mammalian genes.

- Yeast appears to have much the same ability as mammals to add carbohydrate groups to newly synthesized proteins, something bacteria cannot do; the carbohydrate groups may be essential for the activity of certain proteins.

- From at least the time that Bacchus sprang from the thigh of Zeus, there has been an accumulation of knowledge of the large scale culture of yeast. By comparison, bacterial culturing is still in its infancy, which perhaps explains the continuing problem of "phage-out", the disaster that befalls a bacterial culture when it is attacked by a bacterial phage or virus.

Peter Newmark

Thatcher applauds genes

Britain's Prime Minister, Mrs Margaret Thatcher, was told last week that the public stood to gain an annual return of 157 per cent on its investment in basic research into the applications of recombinant DNA techniques through cost reductions in pharmaceutical and other commercial products. In comparison, the heavy investment in research and development, as well as production facilities, meant that companies might on average expect a more modest return of about 15 per cent.

Mrs Thatcher was given these figures by Dr Leslie Glick, president of the Rockville-based company Genex Corporation, when she visited the company's laboratories during her stay in Washington. Dr Glick said that the company had calculated that in 20 years time, \$40,000 million out of \$125,000 million in worldwide sales of a variety of products would be obtained by using genetically engineered microorganisms. Out of the total, \$140,000 million would be obtained from microorganisms making products that already exist, and the remaining \$26,000 from products which are not now economical to make.

These figures would translate into earnings of \$3,500 million for corporations after 20 years, half coming from petrochemicals being made biologically. Total private investment in plant, equipment and research and development would total \$24,000 million; in contrast, public support for basic research was likely to be at the most \$3,000 million, whereas the public would benefit by \$2,400 million in savings and \$2,300 million in taxes on products sold.

Mrs Thatcher said that she "totally agreed" with Dr Glick's reservations about attempting to promote genetic engineering technology through government investment rather than private entrepreneurship. Individual investors were willing to risk something if they stood to gain a great deal, she said, and it was also exciting — "better than betting on horses".

Darwin in California

No verdict

Washington

Biology teachers in the United States are breathing a little more easily this week. Last Friday, a California judge ruled that the teaching of evolution in schools does not necessarily undermine the religious beliefs of those who accept the biblical description of human origins. The claim has been made by Mr Kelly Segraves, director of the Creation Science Research Center in San Diego, in a suit brought against the state's Board of Education on behalf of his son. But, after broad philosophical questions about conflict

between the Darwinian and "creationist" interpretations of biology had been reduced to the narrower question of whether evolution is being taught too dogmatically in Californian schools, the judge ruled that he found little wrong with the state's present practices.

The case had excited wide interest and had been widely talked about as a successor to the Scopes "monkey trial" in Tennessee in 1925. Last week, the creationists swept in like lions, charging that the teaching of Darwinian evolution in schools was an offence to their religious beliefs, and thus a direct infringement of the First Amendment, which guarantees freedom of religion, but they left relatively tamed, arguing merely for changes in the guidelines which the state provides for public school biology teachers on how evolution should be taught. Their demands that alternative theories about human origins should be discussed in biology classes were dropped.

The turning point came on the second day of the trial, after the plaintiffs' attorney had claimed that there was a direct conflict between his clients' beliefs and the state's approach to teaching evolution, exemplified by the current guidelines, which state that evolution "has been going on for so long that it has produced all the groups and kinds of plants and animals now living" as evidence that his clients' religious rights were being infringed.

The attorney, Mr Robert Turner, then surprised the court by withdrawing the original complaint that the state was trying to establish the teaching of a secular religion through school biology teaching. Judge Irving H. Perluss accordingly ruled out debate on the scientific beliefs of Mr Segraves and therefore any attempt to demonstrate the relative merits of creation theories and Darwinism. As a consequence, the defendants were not able, as they had planned, to call the twenty scientists recruited as witnesses for Darwinism by the National Association of Biology Teachers. They included Nobel Laureate Arthur Kornberg and astronomer Carl Sagan.

Another issue ruled out by the judge concerned school teaching practices. On the second day, the plaintiff's son, 13-year-old Kasy Segraves, said he had been taught that man evolved from the apes "as fact". The defendants sought to call his teacher as a witness, both to repeat a denial she had made in the press and to point out that the discussion had taken place in a social science, not a biology class. Judge Perluss refused to allow this line of argument, invoking a state law which bans evidence that would prejudice a case to a greater degree than the value of the evidence.

On the final day of the trial, a member of the state's Board of Education, Mrs Marian Drinker, testified that for the past eight years — and following earlier criticism of approved school textbooks by fundamentalist religious groups — the

state had adopted a policy that Darwinian evolution should be taught as theory and not dogma. The board had at the time issued a statement requiring that in school texts "dogmatism should be changed to conditional statements where speculation is offered for origins", and that they "should emphasize 'how' and not the ultimate cause for origins". In his judgement, the judge directed that the Department of Education should recirculate a memorandum containing this policy statement to all schools in California, to textbook publishers and to anybody else who had received the guidelines in the past or would in the future.

Even this was interpreted as a partial victory by the creationists, who argued that it represented "the first move in a chess game", and vowed to continue their fight against the teaching of evolution in schools. The state authorities and their scientific advisers, however, said that the verdict was a complete vindication of the way that evolutionary theory is at present taught in biology classes, pointing out that the judge in his summing up had stressed the importance of not mixing religion and science.

David Dickson

European space agency

Sixes and sevens

The plans of Eric Quistgaard, the new director-general of the European Space Agency (ESA), have thrown the 11 member states into hot dispute over the agency's future. Three months after the plans were announced, the delegates to the council are no nearer agreeing a draft that would lay down the agency's structure and aims for the next 10 years. The fear that ESA may collapse if member states cannot reconcile their individual ambitions may be realized if agreement is not reached within a few months.

ESA suffers from all the troubles of multi-national, multi-purpose agencies. The latest crisis has been brewing for the past two years as the agency's two largest projects, the Ariane launcher and Spacelab, have been drawing to a close and no new projects of comparable size have been agreed to replace them. Quistgaard's solution, to reduce the agency's total budget from about £480 million to £285 million a year by 1982, has been agreed by most member states. But his suggestions as to how the reduced budget should be spent have opened up old wounds.

West Germany is the only country so far to support Quistgaard's plan to increase the mandatory science budget by 50 per cent over the next decade. Although all are thought to support the notion that the European scientific community should support more space science, few are willing to commit the funds which they would prefer to spend as they wish on voluntary programmes.

The reduced budget will mean that some

selectivity is necessary and the latest dispute seems to be over which programmes the agency should concentrate on. The principal candidates are telecommunications, space transportation systems and Earth resources, with Spacelab and rocket-borne micro-gravity experiments a minor option.

The option standing the greatest chance of success is the Earth resources project. Many member states agree that Europe should be more active in the field, and a decision is due later this year on whether to go ahead with ERS 1, ESA's first remote sensing satellite for oceanographic studies. A more controversial plan for future Earth resources satellites is also to be planned.

But the greatest dispute is over what should be done about telecommunications, supported by Britain, Italy and the smaller countries, and space transportation systems, supported by France. France and Germany, by far the biggest space spenders inside and outside the agency, opted out of ESA's Large Communications Satellite (L-sat) when they decided to build jointly a large satellite for direct broadcasting. Both countries would prefer telecommunications satellites, already showing commercial potential, to be built outside ESA which is primarily a research and development agency.

The future of L-sat will probably be assured later in the year when Britain, Italy and the smaller nations give it their approval. The question then will be whether to build the next generation of advanced telecommunications satellites within the agency or within French and German industry. The issue is particularly sensitive for Britain, which has recently

decided to build up its telecommunications industry in the hope that ESA could provide more support.

The dispute has driven the ESA members into two camps — the smaller countries led by Britain, and France and Germany who put up most of the money for Ariane and Spacelab. It is feared that France and Germany could decide to contribute even less to the agency than M. Quistgaard's plan envisages, preferring instead to put more money into their own well-developed space industries. That could weaken the agency to the point where it could not afford to run its own basic support facilities. As yet, however, according to one official, nothing is certain and anything is possible. The next few months should determine the future.

Judy Redfearn

Chemical warfare

Budget increase

Washington

The first signs that President Reagan will implement his election promise to produce chemical weapons came last week in the Defense Department's budget proposals. The Pentagon has confirmed that the \$222,200 million budget for the next fiscal year will include \$20 million for construction work on a plant at the Pine Bluffs arsenal in Arkansas for producing binary chemical weapons.

Provisional approval for the building of the plant was given by Congress last year, but since then there has been intense lobbying by church groups and other protesters, while the previous Defense Secretary, Harold Brown, was lukewarm on the project. This is why funds for the first stage of its construction were not included in President Carter's budget submitted to Congress early in January.

Pentagon officials are at pains to emphasize that the request for \$20 million does not necessarily mean that Mr Reagan plans to end the moratorium on the production of chemical weapons begun by President Nixon in 1970. But they think that the new president is "leaning in that direction" and that his second budget a year from now would ask for a further \$140 million to build full-scale manufacturing facilities.

Throughout the past decade, the US Army has kept up the pressure for reduced production, arguing that even if only a small stockpile of chemical weapons were required for deterrence, binary weapons could be stored more safely than those in which active chemical agents are present from the start. For some time, the army has been embarrassed by the leakage of nerve gas from "Weteye" bombs stored in Colorado.

Congress refused to fund the construction of binary weapons in 1975 and 1976, with the result that the US Army's interest in going ahead was dis-

couraged by the White House in succeeding years. There is, however, evidence that President Carter may have relaxed his previous opposition to chemical weapons shortly before leaving office. The Arms Control and Disarmament Agency said, when the president's last budget was published in January, that the Administration had agreed to an increase of chemical weapons activity within the Defense Department. The agency estimated that expenditure would grow from \$2 million in the fiscal year 1981 to \$14.4 million in 1982.

Among the projects foreseen in the Carter budget were the experimental development of new binary weapons (\$2.2 million) and the development of a binary spray bomb in which all three services had an interest. Production of this weapon, called "Bigeye", will require both presidential and congressional approval, as will the production of the 8-inch and 155-mm artillery projectiles now under development. One of the issues that will concern both the White House and Congress is the effect of a decision to resume production on the negotiations with the Soviet Union on a chemical disarmament treaty, which has been discussed in Geneva since August 1976.

The Defense Department also asked Congress in January for \$44 million for "chemical defense science and technology" projects in 1982. This is considerably higher than in previous years, and follows recommendations about necessary improvements to protective equipment, warning devices and medical treatment made last summer in a study conducted by the Defense Science Board.

David Dickson

One step to LEP

The UK Advisory Board for the Research Councils has recommended to Secretary of State Mark Carlisle that Britain should support LEP — the SF 960 million electron-positron collider planned for the European subnuclear physics laboratory, CERN, near Geneva. If Mr Carlisle accepts the board's advice, British delegates at CERN's June council meeting will vote for LEP to be included as an integral part of CERN's future programme, not as a separate project as some factions — particularly in Sweden — would like.

Meanwhile, LEP has shrunk — if not in cost, in size. A new design has LEP only 27 km in circumference, so only 8 km of the accelerator tunnel will be under the French Jura Mountains to the north of the CERN site. In this way LEP will avoid a region of fractured and unstable rock — the "trias" — which was threatening substantial increases in civil engineering costs. Improvements in machine design will enable the new LEP to reach the same energy without any increase in power consumption.

Robert Walgate

Allen's successor

Professor John Kingman is to be the next chairman of the Science Research Council. He will take over from Sir Geoffrey Allen, who is returning to his old research group at Imperial College, on 1 October.

Professor Kingman, who will be leaving his niche as professor of mathematics at the University of Oxford, is an unusual choice, having had no experience of experimental science. His work has been in mathematics and statistics. Since 1979, however, he has been chairman of the council's science board. The board's zeal for awarding research grants was largely responsible for overspending by the council to the tune of £10 million in 1979–80.

The Department of Education and Science, which made the appointment, had apparently hoped to boost the council's image as a supporter of biological sciences by appointing a biologist, but no willing candidates could be found.

Judy Redfearn

Chemical industry

ICI off course

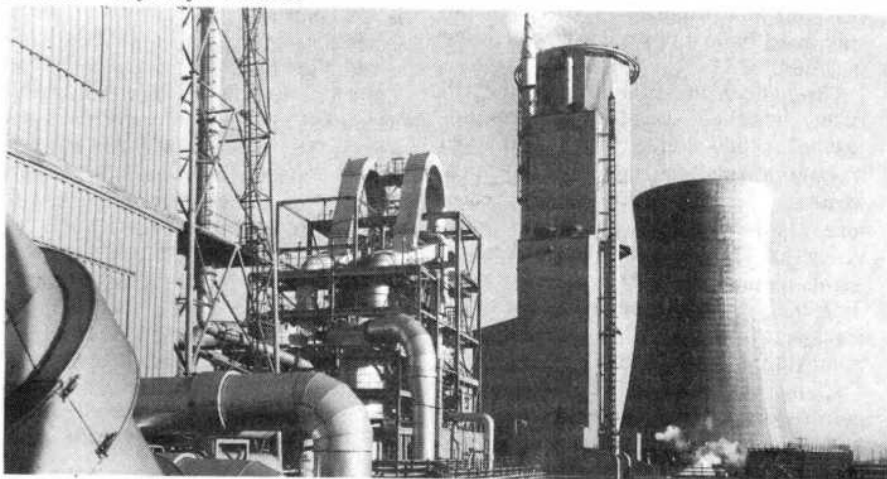
Britain's largest chemical company — Imperial Chemical Industries Limited (ICI) — is turning to biology to seek a long-term solution to its present financial difficulties. Speaking after the company's announcement of its worst share dividend since 1938, ICI's research director, Charles Reece, said that research priorities would now be in biology-related disciplines — coupled with efforts in composite materials and the by-product chemistry of new transport fuels.

A third of ICI's £200 million annual research and development budget is already linked to biology. Most of this goes into pharmaceuticals, forming the largest single component (18 per cent) of ICI's research. ICI also has the world's largest single-cell protein fermenter and has used genetic engineering to modify the protein-producing organism *Methylophilus methylotrophus* to improve its carbon efficiency (see *Nature* 287, 396; 1980). The modified organism is not in production yet, but only "technicalities" delay its use.

Reece would not say what products ICI would make through genetic engineering except "the obvious ones": proteins for human, animal and plant health. The company is paying special attention to plants — too little is yet known of plant molecular genetics to expect a quick return, but since it is an area of ignorance, it is also likely to be an area of opportunity.

But this is long-term thinking. In the short run, ICI is not about to jump on the bandwagon of small-volume, higher added value products recommended by the pundits of the chemical industry — unlike, say, BP Chemicals, which last week announced a £4 million investment in a new plant at Hythe to produce hydroxyalkyl-acrylates for speciality paints and other uses. Speciality chemicals, the argument goes, are — paradoxically — bought by many small consumers, whereas bulk chemicals are bought only by one or two large buyers. So if the bulk buyers go down, so does the supplier.

Pruteen — no food for the birds



ICI is not impressed by this generalist argument. The company is more concerned to estimate the precise needs of existing customers, who in a recession are being more exacting in the use of their money. This means that research and development in ICI should be related to those needs, but the difficulty is to distinguish long-term structural changes in needs from the effects of the recession, says Reece. One structural change might be the price of steel, with its high energy intensity; this is forcing consideration of substitute materials, particularly composites (rather than new polymers).

Another structural change would be the high cost of oil causing shifts to new fuels — such as syncrude from coal — which will have new chemistries and new spin-off feedstocks for the chemical industry.

So research effort should be concentrated on "C-1 and C-2" chemistry, the chemistry of molecules with low numbers of carbon atoms as potential feedstocks. This would also provide flexibility: whatever the principal transportation fuel source turns out to be, its products could be cracked down to low-carbon molecules to begin higher synthesis.

Back in the factory, production processes could be optimized to the new economics — emphasizing thermodynamic balance for energy or feedstock efficiency, for example, or reducing the capital cost of equipment to allow product flexibility.

Tight environmental regulations are strangling new products, Reece believes. "Try to get water through regulation now — you'd fail." There is an element of double standards for new and existing products. But the regulations, seen in another light, can provide competitive research opportunities: to improve the selectivity of a pesticide, for example, so it might pass regulation more quickly.

But, overall, research in ICI is being seen as just another part of the business, which must tighten its belt and show its economic effectiveness like all other divisions. The days when ICI's corporate laboratory in Runcorn, Cheshire, was called "ICI's university" are over.

Robert Walgate

Indian oil exploration

Partners sought

Lucknow

Faced with the urgent need to find more oil, and the accompanying huge capital risk, the Indian government has decided to invite foreign oil companies to conduct oil exploration. A short list of 34 companies has been drawn up and contracts will be let on a production-sharing basis only to companies agreeing to sell their share of crude to India at a fair market price.

The government is offering exploration rights on 32 blocks — 15 onshore and 17 offshore — a total of about 0.66 million km². The public sector Oil and Natural Gas Commission will represent the government as a "non-operating" party in all exploration contracts and will be given full decision-making rights on any discovery.

At present, oil exploration is carried out by the commission and Oil India Limited. The commission has discovered 585 million tonnes of recoverable oil and gas reserves, whereas Oil India has found only 47 million tonnes. According to a report jointly prepared by Soviet and commission experts, however, there are reserves of 12,700 million tonnes of oil, of which 8,700 million tonnes are in offshore and 4,000 million tonnes in onshore areas.

Commercial reserves were first discovered in the early 1920s in the Assam region, and in the mid-1970s crude was commercially produced in Assam and Gujarat states. Offshore exploration began in the early 1960s, and in the mid-1960s reconnaissance surveys by a Soviet seismic ship, Akademik Arkhangeliskey, delineated certain structures including the Bombay High. Oil was discovered there in 1974 and production began in 1976. By last year, 88 exploratory and assessment wells had been drilled, of which 31 had proved to be dry. Other commercial discoveries include the offshore structures Ratnagiri, North Bassein, South Bassein and Tapti.

The Indian Planning Commission has allocated 30,000 million rupees (£1 = R17.90) for oil exploration during the sixth plan period 1980–85, which includes R18,370 million for offshore activity. So far, offshore exploration has cost R7,250 million with total discoveries of 250 million tonnes of oil and 270,000 million m³ of gas — a potential value of R50,000 million in crude alone.

India spends a sizeable amount of foreign exchange earnings on oil imports. During 1981, indigenous production will be 14–15 million tonnes, less than half domestic demand of 31 million tonnes, so R60,000 million will be needed to fill the gap. By 1985, the demand is expected to reach about 47 million tonnes, with indigenous production reaching only 21 million tonnes.

The Bombay High development is being financed by the World Bank, which recently offered a second instalment of

\$400 million, having provided \$150 million in 1977. During their recent conference in Vienna, OPEC finance ministers also approved an interest-free loan of \$30 million for the project. The present plan is that production capacity will rise to 12 million tonnes per annum.

The collaboration now sought with foreign companies is not new to India. In the 1970s, for example, the government invited two American and one Canadian group of companies to explore. They left, however, after finding their wells dry. India has also been collaborating with other countries in contract surveys and recovery or increased production technologies.

Zaka Imam

UK pharmaceuticals

Bad risks

The problems of getting new drugs onto the market have been brought home to two British pharmaceutical companies in the past few weeks. Fisons, after many years of costly development, have had to abandon the antiasthmatic drug Proxicromil in the face of evidence of ulcerogenic and carcinogenic activity in rats. And Beecham's new antibiotic combination, Augmentin, has received a cool reception from the government's Committee on the Safety of Medicines.

In Proxicromil, Fisons thought they had found a successor to their major money-maker Intal (cromolyn). During the past 10 years, Intal has proved to be one of the most effective inhaled antiasthma drugs. But Fisons' patent on Intal lapses in 1982, and other manufacturers will be able to produce alternative forms of cromolyn.

Since 1973, the company has spent between £12 and £15 million on developing Proxicromil as an orally administered alternative. Clinical trials had been very encouraging, and it looked as if Fisons had succeeded where others had failed and had found a drug which could be taken in the high doses needed by the oral route without having unacceptable effects away from the site of action.

But late in the day, only last October, long-term animal toxicity tests began to show that Proxicromil induced ulcers and renal tumour. Fisons acted quickly and halted further work on the drug despite the massive investment already made. Proxicromil had been expected to bring the company earnings of about £10 million a year, and the immediate reaction on the stock market was to reduce the market value of the company's shares by £10 million.

Trouble in the drugs division comes at an awkward time for Fisons. Four fertilizer plants are to be closed, while the company is also suffering from the general slump in the heavy chemicals industry. Last year, Fisons made a net loss of £16.8 million compared with net profits of £12.1 million the year before. Other cost-cutting moves

include closure of the London office.

Beecham's immediate problems are not so grave. But the Committee on the Safety of Medicines, which is technically advisory to the Department of Health, is quibbling with the company over the indications which should be listed for its new antibiotic Augmentin.

This is a combination of the semi-synthetic penicillin amoxicillin with clavulanic acid. The clavulanic acid is there to block the action of β -lactamase enzymes produced by penicillin-resistant organisms and so protect the active amoxicillin.

The committee is broadly suspicious of any combination of drugs, considering that general use might cause problems of multiple drug resistance in the long term. It seeks to restrict the use of Augmentin to only the severest of infections in adults and has suggested it be used only for infections with Gram-negative bacteria, mostly urinary tract infections and bacterial gastroenteritis. Beecham had in mind a wider range of applications. What happens now is that the company will resubmit its views to the committee, pointing out the benefits to be had by a wide use of Augmentin. And sooner or later a final list of indications will emerge.

Charles Wenz

Telecommunications

Bell's bother

Washington

The future of one of the world's largest private research institutions, the Bell Laboratories, once again hangs in the balance following the failure of its parent body, the American Telephone and Telegraph Company (AT&T), to reach a settlement with the US Justice Department on anti-trust charges.

The trial of the Justice Department's suit against AT&T opened on Wednesday, 4 March in New York. The lawyers are settling in for what may be a two-year spell. Last month, a day after the suit had come to court, the case was adjourned when both sides informed Judge Harold H. Greene that the "framework" for a consent decree settlement had been reached. The trial was deferred to allow the details of the settlement to be worked out providing the ground rules under which the company will be allowed to operate in new areas of telecommunications, such as information processing. A fortnight ago, however, negotiations over the terms of the settlement broke down.

The suit filed by the government alleges that AT&T had abused its monopolistic power over the US telephone system in the 1960s and 1970s, and required the company to divest itself of various parts. This would have included not only 23 operating telephone companies and an equipment manufacturer, but also part of the Bell Labs, where much of the world's basic research into telecommunications is

carried out. Although no formal announcement was made by either side about the rumoured settlement, it was generally believed that the laboratories would not have been affected.

With the breakdown of the negotiations, however, there is no longer any assurance that the laboratories will remain untouched, and it seems likely that the final settlement could be less favourable to the company.

A major uncertainty at present is the likely attitude of the new Administration, whose refusal to approve the framework negotiated under the Carter Administration has been cited as one of the reasons for the delay in reaching an agreement.

Last week, President Reagan announced the nomination of Stanford law professor William F. Baxter as Assistant Attorney General in the Department of Justice responsible for anti-trust affairs. Mr Baxter has previously spoken out against large divestiture cases, and the Office of Management and Budget has recently proposed terminating the anti-trust activities of the Federal Trade Commission, both factors which were thought to be in AT&T's favour.

However, lobbyists for AT&T's competitors cite a paper in the American Bar Association's *Antitrust Journal* written by Mr Baxter in 1977 in which he argues that divestiture may be the appropriate remedy in cases involving regulated monopolies.

The resumption of the trial has also provided an excuse for renewed attempts on Capitol Hill to develop legislation aimed at breaking up AT&T, on the grounds that the future control of telecommunications technology is too important to be left to regulatory commissions or to the courts.

Last week, Timothy Wirth, the new chairman of the House subcommittee on telecommunications, consumer protection and finance, said that the House was prepared to step in with legislation to increase competition in the telecommunications industry. Members of the Senate Commerce Committee are also considering draft legislation.

Last year, the House committee passed legislation which would have required the company merely to undergo an internal restructuring, setting up a subsidiary to carry out the research and development of competitive services. Such legislation had long been sought by AT&T, but was opposed by its competitors on the grounds that it would be pre-empting the Justice Department suit.

The latest challenge to AT&T comes from a group of shareholders, who are proposing at the company's annual meeting next month that it change its name to either American Telephone and Technology Inc. or American Telephone and Telecommunications Inc. to reflect the new areas of operation that the company is keen to enter. Both changes are opposed by the company itself.

David Dickson

CORRESPONDENCE

New IQ test?

SIR — Mackintosh¹ finds my theory of intelligence “distinctly simple” but allows that the findings of the inspection-time (IT) studies are “possibly very important”. I largely agree with him on both counts. No one who has steeped himself in modern “cognitive psychology” — with its legions of black boxes, “executive programmes” and “subroutines”, resembling phrenology in all but the apparent infinity of the number of mental mechanisms that are now conjectured — can readily accept that intelligence is truly general or unitary. As Mackintosh writes, “there might be a variety of independent traits, which happened to correlate with one another in the general population.” Such a lofty invocation of chance “happenings” is commonly preferred by cognitivists to any systematic endeavour to explain why so many measures of mental ability correlate positively as they do.

Professor Mackintosh writes: “A correlation between IQ and inspection time does not mean that IQ is mental speed, nor even that speed is one (of several) causes of IQ.” I quite agree with him on the first point — if only for the reason that it is important to distinguish between the hypothesized “mental speed” on the one hand and its natural, developmental products (in conventional IQ) on the other. But Mackintosh offers no support for his suggestion that the causal relationship between speed and IQ “might even go in the reverse direction”. It is easy, and indeed very traditional to suggest that a higher IQ gives a person many advantages. But why should one of the largest of these advantages — larger than those for educational attainment, income or social prestige — be in judging briefly presented line-lengths? Again, if IQ were causal to IT, it should be noted that one of the studies which I have reported (by Brenda Hosie) found a strong correlation (−0.78) between IQ and IT in children who were only four years old. This hardly looks like a straightforward case of a high IQ providing numerous, small and cumulative advantages over a long period of development.

I do not claim to have “proved” that mental speed is psychologically and ontogenetically basic to IQ-differences. Indeed, even when it becomes possible to manipulate inspection times experimentally (by drugs, for example), mental speed differences may not be able to account for at least that 20 per cent of natural IQ variance that is widely agreed to be “environmental” in origin. However, I do suppose that scholars of cognitivist and environmentalist persuasions will need great imaginativeness to formulate explanations of the IQ–IT relationship that do not involve the concept of a unitary, underlying trait for which *g* is the time-honoured name and psychometric indicator, and for which “mental speed” will prove a convenient psychological short-hand.

There are numerous potential applications of the finding that IT (and, as Jensen and the Hendricksons find, “choice reaction time” and the “average evoked cortical potential”) can apparently serve to index intelligence. It should at last be possible to operationalize the

concept of intelligence in studies of young human infants and of other species; and this may lead to rapid advances and changes in the understanding of how intelligence is controlled and of how it develops. But Mackintosh doubts that any results obtained with different human racial groups could influence cherished convictions. I fail to understand why. My own view — outlined in a forthcoming paper — happens to be that conventional measures of fluid intelligence are not quite such pure measures of *g* as Jensen loyally maintains: this is principally because some of them are a little contaminated by *k:m* (or “spatial abilities”). I, for one, would feel obliged to modify this view if IT indices gave Afro-Americans the same degree of disadvantage that they have on conventional measures of IQ. Whether “the critics of IQ” could ever be persuaded to change their views as a result of such studies is not the question. So long as the IQ–IT relationship proves generally replicable, new evidence about IT in non-WASP subjects should have appropriate effects on the scientific views of all reasonable people.

C.R. BRAND

Department of Psychology,
University of Edinburgh, Edinburgh, UK

1. Mackintosh, N.J. *Nature* **289**, 529–530 (1981).
2. Brand, C.R. & Deary, in *Models of Intelligence* (ed. Eysenck, H.J.) (in preparation).

Darwin's survival

SIR — Your leading article “Darwin's death in South Kensington” (*Nature* 26 February, p.735) illustrates “the rot at the museum” by quoting a passage from our 1978 *Guide*. How odd, for that passage was drafted (by me) as a conscious paraphrase of the part of Chapter 13 in *The Origin of Species* in which Darwin discussed the relation between his theory and systematics. “Groups-within-groups classification”, which you take to be a “popular euphemism for cladism” and its attendant heresies, is not hidden propaganda but a contraction of Darwin's words — “the grand fact in natural history of the subordination of group under group.”

The “weasel words” which so incense you are “If the theory of evolution is true.” I have tried replacing them by your own criterion of truth: “If the theory of evolution is not an open question among serious biologists, the features used to classify species in groups . . . were acquired by the common ancestor of the group.” It does not read well.

Your readers may try the substitution in the equivalent passage from Darwin: “on the view that the natural system is founded on descent with modification . . . the characters which naturalists consider as showing true affinity between any two or more species, are those which have been inherited from a common parent.” And your readers can answer for themselves your question “what purpose except general confusion can be served by these weasel words?” The reader may also be able to judge whether the rot is to be found here or in Little Essex Street.

COLIN PATTERSON

British Museum (Natural History),
London SW7, UK

SIR — As working biologists at the British Museum (Natural History) we were astonished to read your editorial “Darwin's death in South Kensington” (*Nature* 26 February, p.735). How is it that a journal such as yours that is devoted to science and its practice can advocate that theory be presented as fact? This is the stuff of prejudice, not science, and as scientists our basic concern is to keep an open mind on the unknowable. Surely it should not be otherwise?

You suggest that most of us would rather lose our right hands than begin a sentence with the phrase “If the theory of evolution is true . . .” Are we to take it that evolution is a fact, proven to the limits of scientific rigour? If that is the inference then we must disagree most strongly. We have no absolute proof of the theory of evolution. What we do have is overwhelming circumstantial evidence in favour of it and as yet no better alternative. But the theory of evolution would be abandoned tomorrow if a better theory appeared.

Charles Darwin died nearly a century ago and is honoured at South Kensington as a great man of science. It does neither him nor science any service to misrepresent the status of his work.

H. W. BALL, A. GRAY, L.A. MOUND,
J.F.M. CANNON, C.J. HUMPHRIES,
H.M. PLATT, G.C.S. CLARKE, P.W. JAMES,
J.H. PRICE, R.F. EASTWOOD, A.C. JERMY,
N.K. B. ROBSON, P.L. FOREY, D.M. JOHN,
C.B. STRINGER, J.D. GEORGE, S.W. JONES,
D.A. SUTTON, M.GIBBY, R.S. MILES,
R.I. VANE-WRIGHT AND P.J.P. WHITEHEAD
Departments of Botany, Entomology,
Palaeontology, Public Services and
Zoology, British Museum (Natural History),
London SW7, UK

Conservation sites

SIR — It can hardly be doubted that the number of species alive today is much smaller than the number of extinct species, nor that the future, unless it be preternaturally terminated, will see a proliferation of types of life exceeding all which have gone before. From this I conclude that extinction is the normal if not the necessary consummation of any species. If this is so, to attribute value to any species simply because it happens to co-exist with an observer, and still more, to attribute greater value to a species which, by its rarity, demonstrates its biological deficiency, reveals in that observer either or both a level of prejudice and lack of reason which gives cause to doubt his judgement on the matter of conservation even though he be a professional. Yet this is the essence of the action of the Nature Conservation Council in designating “sites of special scientific interest”. Were these sites to be set up with no loss of freedom to anyone, and more importantly, no cost to any individual, their institution might be treated as innocent caprice. The real consequence of the foolish advice given by the Nature Conservation Council is that their political masters impoverish the owners of the land involved, in the belief that they are preserving a national or even international treasure. Dr Goode of the

Continued on page 173

NEWS AND VIEWS

The intergalactic medium

from Joseph Silk

To many astronomers, the search for intergalactic matter resembles the quest for the holy grail. Theory tells us unhesitatingly that the remnants of the huge clouds out of which galaxies formed should still be in intergalactic space. The process of galaxy formation, like practically everything else in astronomy, was inefficient. The vestiges of protogalaxies, or embryonic galaxies, should pervade the Universe as wispy filaments or clouds, and be embedded in a uniform intergalactic medium (IGM) from which every structure we observe today must once have condensed. Theory tells us, and here it develops philosophical undertones, that the Universe was once very simple. A homogeneous and isotropically expanding substrate is the stuff out of which our cosmological models are made, as are the galaxies. Small fluctuations and slight deviations from perfect uniformity were amplified by the inexorable process of gravitational instability, eventually growing into protogalactic gas clouds that collapsed and fragmented into vast systems of stars.

How much matter is left behind is an open question. It is very likely that only a small fraction of the matter in the Universe has formed the visible regions of galaxies. Observations of galaxy rotation and the dynamics of galaxy groups and clusters have indicated the existence of a considerable amount of dark matter. Unseen by any astronomer, this matter nevertheless dominates the gravitational potential in the haloes of galaxies and in bound systems of galaxies. It is a component of the Universe, perhaps the dominant component, that has failed to collapse and make stars in the same manner as the inner regions of galaxies. The dark matter does not consist of gas, otherwise it would be visible. But what it is, we cannot tell. Speculations abound, the current favourite being massive neutrinos.

The luminous regions of galaxies consist overwhelmingly of stars, and these have formed from clouds of gas. We see this process of star formation continuing today in interstellar molecular clouds. The Milky

Way galaxy is a mature system, most of its stars having formed long ago from a much larger gas cloud. The remnants of clouds like this, and the gaseous medium out of which the cloud itself condensed, constitute the elusive IGM.

The search for traces of the IGM has been long and often fruitless. Reputations have been made and broken, often only for what has finally proved to be a new, perhaps stronger, upper limit on what might or might not be present in intergalactic space. Radio, optical, space ultraviolet, X-ray and γ -ray astronomers have all confronted the challenge, and all have failed to find the elusive, uniform IGM.

One early attempt utilized the 21-cm absorption line of atomic hydrogen to search for neutral intergalactic gas along the line of sight to the radio galaxy Cygnus A. No feature was found, and an upper limit was set.

The most stringent limit on the presence of a neutral intergalactic medium came with the discovery of quasars. Sufficiently distant that the cosmological redshift brought the Lyman α line of hydrogen into the visible part of the spectrum, quasar spectra were searched for a possible absorption 'trough' due to intergalactic neutral hydrogen. Failure to detect such a feature meant that, in the region of the Universe that was probed, less than one part in a million of an IGM of density sufficient to close the Universe can exist in neutral form. Subsequent space observations have extended the search for a Lyman α absorption trough into the ultraviolet region of the spectrum in the nearby quasar 3C 273, and confirmed that the IGM, if it exists, must be highly ionized and hotter than 10^6 K.

Such a hot IGM would lead to copious X-ray emission and the discovery in the earliest days of X-ray astronomy of a diffuse background of X-ray emission

produced some evidence for it. However, recent satellite observations with the Einstein observatory have revealed that quasars can account for a substantial fraction (over 50 per cent) of the X-ray background. At the same time, evidence for hot intergalactic matter has been found in X-ray studies of galaxy clusters. These contain substantial amounts of X-ray-emitting gas that is enriched to near the level of the Sun with certain heavy elements. High-resolution observations of quasar spectra have also revealed the presence of discrete intervening clouds of intergalactic gas. But despite the accumulating evidence for intergalactic matter in localized regions, the hypothesized uniform IGM that is so cherished by cosmologists has still not been found.

A recent paper by Paul Shapiro and John Bahcall (*Astrophys. J.* **241**, 1; 1980) makes an especially intriguing proposal which will extend the search for a Lyman α absorption (which led to a severe limit on a neutral IGM) to the remaining possibility of an ionized IGM. The authors suggest that detection of a hot, ionized IGM may be possible through observations of the X-ray absorption spectra of quasars. Extending earlier calculations by Richard Sherman (*Astrophys. J.* **232**, 1; 1979) they note that quasars, as a class, are sufficiently luminous X-ray sources to be observable at a redshift greater than 2 or 3. If sufficient atoms heavier than hydrogen are present in a hot IGM they will produce absorption troughs in the X-ray region of the spectrum, due to excitation of inner electrons around the heavy atomic nucleus; iron, for example, has a characteristic absorption at a few keV. Broad absorption troughs in the X-ray spectra of distant quasars would therefore be strong evidence for a hot, ionized IGM with a density amounting to a substantial fraction of the closure density and containing a near solar abundance of heavy atoms. Future space experiments with X-ray detectors capable of moderate energy resolution and high sensitivity should be able to perform this measurement within the next decade. □

Joseph Silk is a professor at the Department of Astronomy, University of California at Berkeley.

Directing toxins to cancer cells

from Sjur Olsnes

A CENTRAL PROBLEM in cancer chemotherapy is the lack of selectivity of present anti-cancer drugs. One attempt to improve this selectivity has involved the linking of cytostatic drugs, like daunomycin, adriamycin and methotrexate, to antibodies against tumour-associated surface antigens^{1,2}. The ability of such complexes to kill malignant cells selectively has, so far, been limited. Blythman and his colleagues (see this issue of *Nature*, p.145) are among those who have tried an alternative approach, whereby a toxin, instead of a drug, is linked to the antibody, and have now succeeded in producing a complex which is not only toxic but also highly selective.

The limited success of the drug-antibody complexes may be partly due to the fact that the cytostatic drugs act stoichiometrically and a large number of molecules is required to kill a cell. Even if the antibodies have the required selectivity (and it is not yet clear to what extent the corresponding antigens are tumour specific), they are usually obtained in such low amounts that the number of complexes reaching the tumour cells and binding to them may be very small. Furthermore, if the cytostatic drugs are for some reason released from the antibody, their non-selectivity is restored.

Interest in the preparation of tumour-specific cytostatics has recently increased for two reasons. First, the hybridoma technique permits the production of large amounts of monoclonal antibodies to a variety of cell-surface antigens, including tumour-associated ones. Second, the extreme toxicity of certain bacterial and plant toxins — a single molecule may kill a cell — might compensate for the small amount of complexes accumulating in the tumour.

Diphtheria toxin, abrin, ricin and related toxins all consist of two polypeptides with different functions. A binding polypeptide (B), with specificity for receptors at the cell surface, is linked by a disulphide bond to an enzymatically active polypeptide (A), which enters the cytoplasm and inactivates components of the protein-synthesizing machinery. Clearly, if the B-polypeptide is replaced by an antibody to a tumour-associated antigen, the enzymatically active A-polypeptide might be selectively delivered to the malignant cells. It is also important that the A-polypeptide alone is not cytotoxic.

In the first attempts to link diphtheria toxin to antibodies^{3,4} only marginal selectivity was achieved, but later studies have been much more successful. Thorpe *et al.*^{5,6} linked diphtheria toxin to anti-lymphocyte globulin and to F(ab')₂

fragments derived from it. These complexes were highly toxic to cultured human lymphoblastoid cells which are rather insensitive to diphtheria toxin alone. A chimera of diphtheria toxin fragment A and Fab' fragments of IgG from antiserum against L 1210 cells also showed selective toxicity in cell culture⁷. An interesting toxic conjugate was formed of ricin A-chain and anti-idiotypic antibodies against surface immunoglobulins on a B-cell lymphoma⁸. This procedure may have a more general application as idiotype antigens on the surface of B-cell lymphomas are, in fact, specific for the particular tumour cells.

During the past year four groups have reported experiments linking toxin A-chains by a disulphide bond to monoclonal antibodies against cell-surface antigens. In one case the antibody was against a colorectal tumour-associated antigen⁹, whereas in the others they were against normal surface antigens like IgM and IgD on B lymphocytes⁸ and the differentiation antigen Thy-1.2 on thymocytes¹⁰. In each case the chimaeric toxins were selectively toxic to cultured cells carrying the corresponding antigen.

The chimera or 'immunotoxin' formed by Blythman *et al.*, between ricin A-chain and a monoclonal IgM antibody against the Thy-1.2 antigen, was almost as toxic as native ricin to murine lymphoma cells (WEHI-7) carrying the Thy-1.2 antigen. In contrast, cells lacking the Thy-1.2 antigen tolerated 2,000–6,000 times more immunotoxin although they were fully sensitive to native ricin. The authors also demonstrate that treatment with immunotoxin significantly increased the survival time of mice challenged with WEHI-7 lymphoma cells. Although promising, the protection *in vivo* was less than expected from the cell culture data. There may be several reasons for this discrepancy. As pointed out by the authors, the two test systems used are entirely different — *in vivo* a few surviving cells may grow up to tumours and kill the animal, whereas in the cell culture assay a few surviving cells would hardly be recognized. Furthermore, *in vivo* the access of immunotoxins to the lymphoma cells may be reduced, for example due to the removal of the chimaeras by reticuloendothelial cells¹¹. Clearly, pharmacokinetic studies will be required to design the optimal type of immunotoxin.

Binding moieties other than antibodies have also been exploited. Several groups have made cytotoxic chimaeras by linking toxin A-chains to concanavalin A^{12,13} and other lectins¹⁴, to asialofetuin¹⁵ which is taken up selectively by receptors on hepatocytes, and to monophosphopentamannose¹⁶ which is bound and internalized by receptors on fibroblasts.

Particularly interesting are chimaeras between toxin A-chains and hormones like insulin¹⁷, the β subunit of human chorionic gonadotropin¹⁸ and epidermal growth factor¹⁹. The fact that the hormones are able to direct the toxin A-chain to the cytoplasm and kill the cell may be used to select mutant cells lacking hormone receptors and mutants with defects in the route of internalization. Such mutants may elucidate the functional significance of receptor-mediated endocytosis of polypeptide hormones. Not all hormone receptors may be able to internalize chimaeric toxins, as shown by Chang *et al.*²⁰ who could not demonstrate any toxic effect of a conjugate of diphtheria toxin fragment A and human placental lactogen.

Future work may take advantage of the fact that many plants accumulate considerable amounts of A-chain-like proteins which, as such, are not toxic to cells as they lack a B-chain. One of these proteins, gelonin, which is particularly abundant, became cytotoxic after being linked by a disulphide bond to concanavalin A²¹.

A limitation to the therapeutic use of chimaeric toxins is that, as foreign proteins, they elicit the production of neutralizing antibodies. This can, however, be considerably retarded if the toxins are given together with immunosuppressive cytostatic drugs²². Furthermore, the present availability of a variety of immunologically unrelated A-chains may allow a switch to another chimera as soon as an immune response is detectable. □

1. Levy, R., Hurwitz, E., Maron, R., Arnon, R. & Sela, M. *Cancer Res.* **35**, 1182 (1975).
2. Ghose, T. & Blair, A.H. *J. natn. Cancer Inst.* **61**, 657 (1978).
3. Moolten, F.L. & Cooperband, S.R. *Science* **169**, 68 (1970).
4. Philpott, G.W., Bower, R.J. & Parker, C.W. *Surgery* **73**, 928 (1973).
5. Thorpe, P.E. *et al. Nature* **271**, 752 (1978).
6. Ross, W.C.J. *et al. Eur. J. Biochem.* **104**, 381 (1980).
7. Masuh, Y., Hara, T. & Noguchi, T. *Biochem. biophys. Res. Commun.* **90**, 320 (1979).
8. Kroleck, K.A., Villemez, C., Isakson, P., Uhr, J.W. & Vitetta, E.S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5419 (1980).
9. Gilliland, D.G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4539 (1980).
10. Youle, R.J. & Neville, D.M. *Jr Proc. natn. Acad. Sci. U.S.A.* **77**, 5483 (1980).
11. Refsnes, K. & Munthe-Kaas, A.C. *J. exp. Med.* **143**, 1464 (1976).
12. Gilliland, D.G., Collier, R.J., Moehring, J.M. & Moehring, T.J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5319 (1978).
13. Yamaguchi, T. *et al. J. natn. Cancer Inst.* **62**, 1387 (1979).
14. Uchida, T., Medada, E. & Okada, Y. *J. biol. Chem.* **255**, 6687 (1980).
15. Cawley, D.B., Simpson, D.L. & Herschman, H.R. *Fedn Proc.* **39**, 1798, Abstr. 1015 (1980).
16. Youle, R.J., Murray, G.J. & Neville, D.M. *Jr Proc. natn. Acad. Sci. U.S.A.* **76**, 5559 (1979).
17. Miskimins, K.W. & Schimizu, N. *Biochem. biophys. Res. Commun.* **91**, 143 (1979).
18. Oeltmann, T.N. & Heath, E.C. *J. biol. Chem.* **254**, 1028 (1979).
19. Cawley, D.B., Herschman, H.R., Gilliland, D.G. & Collier, R.J. *Cell* **22**, 563 (1980).
20. Chang, T.-M., Dazord, A. & Neville, D.M. *Jr J. biol. Chem.* **252**, 1515 (1977).
21. Stirpe, F., Olsnes, S. & Pihl, A. *J. biol. Chem.* **255**, 6947 (1980).
22. Moolten, F.L., Capparelli, N.J. & Zajdel, S.H. *J. natn. Cancer Inst.* **55**, 709 (1975).

Sjur Olsnes is at the Norsk Hydro's Institute for Cancer Research, Montebello, Oslo.

The world's worst weeds

from Robert M. May

MUCH has been written about the crop losses caused by insect pests and pathogens (see, for example, *Nature* 267, 669; 1977). Relatively less attention, however, is usually given to the roughly equal losses caused by weeds, which diminish agricultural yields and degrade rangelands and pastures by competing for water, light and nutrients.

Some time ago, in a survey of the problems caused by weeds in developing countries, Holm (*Weed Sci.* 17, 113; 1969) drew up a list of the ten worst weeds in the world. This list was based on guesses about the global losses of major crops, using data from the FAO and from the unfortunately sparse research literature.

Purple nutsedge (*Cyperus rotundus* L.)

Bermuda grass (*Cynodon dactylon* L. Pers.)

Barnyard grass (*Echinochloa crusgalli* L. Beauv.)

Jungle rice (*Echinochloa colonum* L.)

Goosegrass (*Eleusine indica* L. Gaerth.)

Johnson grass (*Sorghum halepense* L. Pers.)

Guinea grass (*Panicum maximum* Jacq.)

Water hyacinth (*Eichhornia crassipes* Mart. Solms)

Cogon grass (*Imperata cylindrica* L. Beauv.)

Lantana (*Lantana camara* L.)

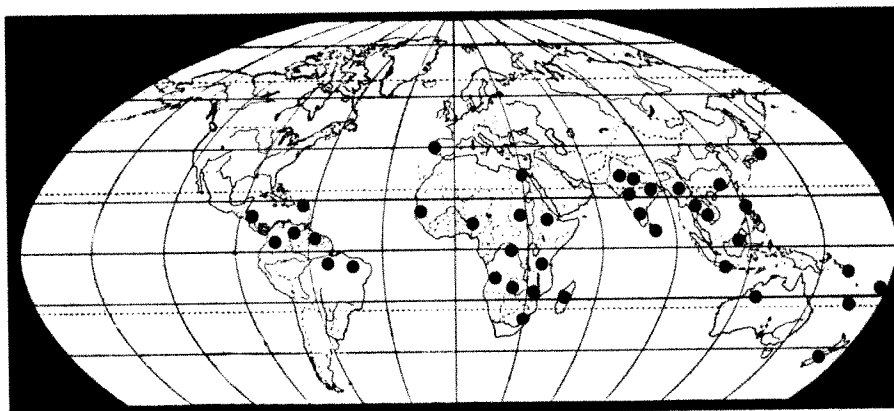
Of these, eight are grasses or sedges and five are perennial grasses.

These villains are typified by the water hyacinth, of which Holm writes: "I have seen it in some of the major rivers of the world and I think it is the most massive, the most terrible and frightening weed problem I have ever known". The figure shows its present world-wide distribution. In developing countries, it chokes waterways and denies riverine communities access to the needed protein from fish in conventional fishing grounds. In the Far East, floods cause great islands of the weed to come crashing through the fences placed in rivers for fish culture. In addition, the weed provides an excellent habitat for many insects that are vectors of diseases of humans and other animals. The water hyacinth was deliberately introduced into the US as an ornamental plant in 1884, as part of a horticultural exhibition. Within 13 years, it had so clogged many waterways as to interfere with navigation, and by 1956 losses in the southern states were estimated at around 45 million dollars per year. In 1976, 15 million dollars were spent on efforts to control water hyacinth in Florida and Louisiana (Cox *Utah Sci.* 38, 59; 1977).

The story of the water hyacinth in the US is typical. Many of the world's top ten,

along with other weeds, have been purposefully introduced into America for one or other reason that seemed sensible at the time. Today, Williams (*Weed Sci.* 28, 300; 1980) estimates that such deliberately introduced weeds cost the agricultural industry millions of dollars each year in reduced yields of crops and livestock, and millions more in control costs. In Williams' words: "Many of these undesirable plants infest crops, others clog navigable or recreational waterways. Still others reduce the productivity and value of grazing lands,

area where halogeton was indigenous. The plant proved salt and drought tolerant, and its spread was not hindered, as it showed promise for forage and firebreaks; the literature of the time knew nothing of its poisonous properties. During the years of depression and World War II, when few weed scientists existed in the US and little money was available for research, halogeton spread rapidly throughout the Intermountain Region. As Williams observes: "By the time sheep losses were traced to halogeton, the plant was beyond



Map showing the distribution of water hyacinth over the warm regions of the world.

or they are poisonous and present a threat to the health of humans and livestock".

Williams asks how best to prevent accidental introduction of weeds, and he gives a tentative sketch of a programme to this end.

For non-poisonous plants, identification of potential pest species should not be too difficult, even though the flora in question ranges from small aquatics to exotic trees. Williams notes that a candidate for introduction should be examined for its capacity to invade crop lands, grazing lands, forests and waterways, paying attention to its methods of dissemination and the longevity of its seeds or other reproductive parts. One should be alert for undesirable characteristics known to be associated with other members of the genus or family to which the species belongs.

For plants that are pests by virtue of being poisonous, the screening process is trickier. Such plants are likely to be considered innocuous in their indigenous setting, or they would not be candidates for introduction; but the plant may, in fact, be poisonous, or may become so under unfamiliar environmental circumstances.

Williams points to halogeton as a classic example of a poisonous weed, whose introduction has been very costly. The plant was accidentally introduced into the western US around 1935, with the first plants probably growing from seed spilled from a train carrying products from an

control, much less eradication. Within 25 years after its discovery, halogeton had engulfed between 4.0 and 4.5 million hectares of desert range in Utah, Nevada, Idaho, Montana and Wyoming, and had invaded Colorado, Oregon and California. Thousands of sheep were killed by halogeton, sometimes 500 to 1,700 in a single day. Even though the plant is widely recognized and largely avoided today, single losses of several hundred sheep are often reported". Today a weed physiologist would recognize that salt-tolerant members of the family Chenopodiaceae often synthesize toxic amounts of soluble oxalates. Analysis of the plant for this compound would be indicated, and the deadly potential of halogeton and its implications for the sheep industry would immediately be apparent.

Williams suggests a combination of several lines of research towards early identification of such poisonous weeds. First, seed inventories should be inspected; a plant may be known to be poisonous, or may be closely related to a known toxic plant. Second, suspected plants should be subjected to chemical analysis for alkaloids, nitrates, nitro compounds, saponins, soluble oxalates, cyanogenetic glycosides, furocoumarins or other exotic toxins indicated for a given plant. Such indications may come from the third line of defense, chemotaxonomy; many plant taxa are related chemically, as well as

Robert M. May is Class of 1877 Professor of Zoology at Princeton University.

0028-0836/81/110085-02\$01.00

© 1981 Macmillan Journals Ltd

morphologically. As mentioned above, such chemotaxonomic considerations would today raise a red flag against halogeton. Another example is the sicklepod milkvetch (*Astragalus falcatus* Lam.), which was introduced around 1920 and has been recommended for range plantings for over 25 years; recognition that many *Astragalus* species contain toxic concentrations of nitro compounds led to identification of sicklepod milkvetch as poisonous, and efforts are now being directed to its eradication. Fourth,

introduced plants that are suspected of being poisonous may be fed to experimental animals, as extracts, dried plants, or residues.

With current pressures to strip-mine for coal, there is much interest in finding plant species that will grow on mine spoils, reproducing rapidly and providing erosion control, and, if possible, grazing for livestock and food and cover for wildlife. The clear need for such plants to be non-poisonous gives special urgency to the programme outlined above. □

Geomagnetic field modelling using satellite data

from David R. Barraclough

ACCURATE DESCRIPTIONS of the Earth's magnetic field are useful for navigational purposes and form the basis for many scientific studies. Such descriptions usually take the form of mathematical models, most often in terms of an expansion in spherical harmonics, and magnetic charts can readily be produced from the models. A major difficulty in producing geomagnetic field models has always been the acquisition of an up-to-date and well distributed set of accurate observations. As a satellite in a polar orbit can survey the entire globe in a matter of days, it was an obvious step to use such a satellite to measure the geomagnetic field.

The most recent satellite survey of the geomagnetic field was performed by the NASA/US Geological Survey satellite Magsat between November 1979 and June 1980. Magsat differed from previous geomagnetic survey satellites in that it was able to measure not only the strength of the field but also its direction.

Since the geomagnetic field is derivable from a scalar potential which satisfies Laplace's equation, it was possible to use the earlier satellite observations of F to refine existing models of the field. Although such models fitted the input F data very well (with a r.m.s. residual of about 10 nT), there was evidence that they did not describe the vector geomagnetic field as accurately (Ben'kova *et al.* *Bull. int. Ass. Geomag. Aeron.* No. 152-163; 1971; Fabiano & Peddie *J. geophys. Res.* **76**, 3816; 1971). These findings led to several theoretical studies of the production of field models from F observations alone.

Hurwitz and Knapp (*J. geophys. Res.* **79**, 3009; 1974) and Barraclough and Nevitt (*Phys. Earth planet. Interiors* **13**, 123; 1976), using simulated data with noise added, verified that good fits to F did not

guarantee a good description of the other field components. In particular, large errors in the vertical component (Z) of the field occurred near the dip-equator (the line of zero inclination). In terms of the coefficients of the spherical harmonic model, those having equal values of m and n were least well determined. These are the sectorial harmonics, which divide the surface of a sphere into segments like those of an orange. It was also shown that the large errors could be removed by using small amounts of vector data (about 10 per cent of the total) covering the equatorial region.

Hurwitz and Knapp (*op. cit.*) and Lowes (*Geophys. J. R. astr. Soc.* **42**, 637; 1975) showed that the root of the problem lay with what they termed the perpendicular error effect. When fitting only a single component of a vector field, the perpendicular components synthesized from the resulting model can have errors associated with them which are considerably larger than the errors in the component fitted. This results from the lack of constraints in the perpendicular directions. In some cases, for example the production of models from Z data only, the perpendicular error effect is of minor importance because of the particular configuration of the geomagnetic field. It turns out, however, that fitting F data is a particularly bad choice. The perpendicular error effect leads to large errors in Z and these are especially troublesome where Z is small, that is, near the dip-equator, as noted above.

Backus (*Q. Jl Mech. appl. Math.* **21**, 195; 1968; *J. geophys. Res.* **75**, 6339; 1970; *Geophys. Res. Lett.* **1**, 21; 1974) investigated whether models derived from F data alone were unique and showed that, under certain conditions, they were not. In practical field modelling using real data these conditions for non-uniqueness are not fulfilled. Stern and Bredekamp (*J. geophys. Res.* **80**, 1776; 1975) have, however, shown that certain series of terms

in the spherical harmonic expansion, which Backus highlighted in his uniqueness studies, are likely to be relatively poorly determined in an analysis of F data alone. These series (which Stern and Bredekamp called 'Backus series') consist of terms which all have the same value of m . Each series starts with a sectional ($n = m$) term and succeeding terms have values of n equal to $m + 2, m + 4, m + 6$, and so on. In agreement with the findings noted above, Stern and Bredekamp found, using simulated data, that the terms with $m = n$ had the largest uncertainty for a given series.

Now that data are available from Magsat, it is possible to study these effects using real data. Stern *et al.* (*Geophys. Res. Lett.* **7**, 941; 1980), in one of the first reports of results from Magsat, have made an initial attempt at this. They have compared two models, one based on vector data from early in the Magsat mission (November 1979) and the other based only on scalar (F) data collected at about the same time. Values of Z computed from the two models differed by up to 2,300 nT at the Earth's surface and by up to 1,500 nT at the satellite's altitude. (For comparison, typical mid-latitude values of Z are of the order of 40,000 nT.) The differences, when plotted on a world map, showed a pattern of six foci, all positioned very near to the dip-equator. When the two models were compared coefficient by coefficient, those in the Backus series described above did indeed show larger differences than the coefficients not belonging to such series. There were suggestions that other effects were present affecting the 'non-Backus' terms and work is in progress in this area.

Thus, the doubts cast on geomagnetic field models based solely on F data seem to be confirmed. Fortunately there are relatively few such models and most of these have been produced for specific purposes related to the scalar data on which they were based. Most producers of models for navigational and general scientific purposes (for example, Barraclough *et al.* *Geophys. J. R. astr. Soc.* **43**, 645; 1975; Barker *et al.* *Geophys. J. R. astr. Soc.* in the press) have been careful to use F observations as only a part of the input data and their models will therefore be free from the deficiencies discussed above. Potential users of field models should be wary in their choice of models. Field models based solely on F data can be used with some confidence if one is interested only in the intensity of the geomagnetic field. If vector information is needed it is best to avoid such models and to use instead one of the models based on component data as well as F data. The International Association of Geomagnetism and Aeronomy, at its next scientific assembly this summer, plans to recommend an International Geomagnetic Reference Field which should provide a set of reliable models describing the geomagnetic field since 1965. □

David R. Barraclough is in the Geomagnetism Unit, Institute of Geological Sciences, Edinburgh.

Hydrothermal activity at mid-ocean ridge axes

from J.M. Edmond

THE existence of hydrothermal activity at oceanic spreading centres has long been suspected but it was only four years ago that the first direct observations¹ were made. During dives of the submersible *Alvin* in Spring 1977, fields of hot springs were discovered at a depth of 2,500 metres on the Galápagos spreading ridge (86°W). The hot spring water had an exit temperature 15°C above that of the surrounding ocean (~2°C) and had produced a thin coating of manganese oxide on the surrounding basalt pillows. Fields of hydrothermally produced mounds composed of nontronites (iron-rich clay minerals) and covered by a pure manganese dioxide crust were also found off-axis in sediments covering 300,000-year old crust².

Systematic work on the chemical composition of the hot springs showed them to be seawater that had reacted with hot basalt at a temperature of about 350°C and then mixed with waters from the ambient ocean³. It was concluded that the mixing had occurred sub-surface and that the observed temperatures in the throats of the vents, and their variations, depended on the extent of this process. Given the highly permeable crust close to the ridge axis it appeared unlikely that the 'neat' high-temperature end-member would ever be directly observed although it seemed likely to be an ore-transporting solution. Since the sulphide-forming elements are removed as the temperature and pH fall during mixing and cause the disseminated mineralization observed in dredged and drilled rocks, it is not possible to 'see through' the mixing zone to the ore metal chemistry of the primary solution⁴.

Most of the major ions — sea salts — are unaffected by this, however. It was therefore possible to estimate the effect of ridge crest hydrothermal activity on the bulk composition of seawater itself. The elemental fluxes are calculated as follows: the thermal budget has been estimated from the observed deficit in the conductive heat flow from young crust relative to that expected from simple plate models. The globally integrated value for the anomaly is approximately 5×10^{19} cal yr⁻¹. If all this heat is removed by hydrothermal convection at an operational temperature of 350°C, then a volume of water equivalent to the whole ocean must circulate through the ridge axis every eight million years. The flux of ³He required to maintain the saturation anomaly at mid-depth in the oceanic water column is roughly 6.5×10^{26} atoms yr⁻¹. The measured ratio of heat to ³He in the hot spring waters is 7.6×10^{-8} cal per atom⁵. If

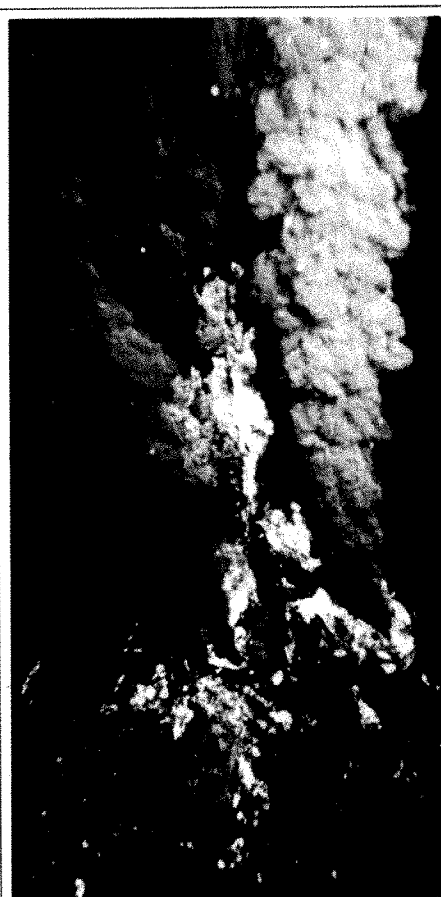
it is assumed that all the ³He is supplied to the ocean by axial hydrothermal vents in the same ratio, the heat flux can be calculated. The value 4.9×10^{19} cal yr⁻¹ is in excellent agreement with the geophysical estimates. It seems plausible, therefore, to compute the fluxes for other elements in an analogous way. Although the circulation rate for water is only about 0.5% of that

input. Lithium, rubidium, calcium and barium are released in amounts comparable to or greater than those delivered from the continents. The manganese input from the axis is sufficient to account for the entire authigenic inventory of the element in deep sea sediments. The contemporary ocean is as volcanogenic as it is fluvial.

Exploration was next extended in areas of known hydrothermal activity in the hope that higher exit temperatures would be found. In 1978 the exciting news was received that the French CYAMEX team had discovered large deposits of iron-zinc-copper sulphides in the axial valley of the East Pacific Rise at 21°N, south of the mouth of the Gulf of California, with their manned diving saucer *Cyana*⁶. While no active hot springs were encountered, this was the first firm indication that something approaching the 'end-member' solutions could survive undiluted to the surface. The planned return to the original Galápagos site in spring of 1979 produced no dramatic developments. The subsequent *Alvin* dives at 21°N did. The French programme had been a reconnaissance survey for a collaborative geophysical programme. Two deep-towed instruments were involved: ANGUS, a sophisticated survey camera, and Deep Tow, a powerful array of geophysical sensors. Ocean-bottom seismometers were deployed in the axial valley, near-bottom gravity and magnetic measurements were made from *Alvin* and active ore-carrying hot springs were found⁷!

'Black smokers' — jets of mineral-laden waters — were observed debouching from constructional features several metres high. These 'chimneys' are formed of sulphides, sulphates, oxides and silicates precipitated as the hot spring waters mix with the ambient seawater. Subsequent dives in the area in the winter of 1979 found that the smokers actually exit as clear, homogeneous solutions with temperatures close to 350°C, as had been previously predicted from the Galápagos data. The major ion chemistry is very similar to that deduced from the low-temperature observations⁸. The end-member had been found. The 350°C waters are completely lacking in magnesium and sulphate, have an *in situ* pH of 4.7 and contain about 2 mmol l⁻¹ of iron and 0.1 mmol l⁻¹ of zinc. The black 'smoke' is composed of sulphides whose precipitation is induced by rapid mixing with the ambient water column. The ore-forming solutions for massive sulphide deposits, the subject of so much speculation, can now be studied directly.

In parallel with the exciting discoveries at Galápagos and 21°N there have been major developments on other fronts.



Picture taken by the research submersible *Alvin* at a depth of about 2,500 m on the crest of the East Pacific Rise in 21°N showing a 350°C hot spring typical of those recently discovered there. The water emerges as a clear homogeneous solution from orifices about 15 cm in diameter at rates of a few metres per second. Rapid mixing with the local ambient seawater induces the precipitation of a variety of minerals from solution. These include sulphides of iron, zinc and copper and several silicates. The 'black smokers' also deposit constructional features — chimneys — of the same materials. These form the conduits emerging from the fields of pillow basalts as seen in the lower part of this picture.

transported by rivers, the much higher salt content of seawater allows for extensive reaction. The composition of the original seawater is radically transformed in the high-temperature zone. The fluxes of dissolved materials into and out of the axis are comparable to those of rivers. Magnesium and sulphate are quantitatively removed, resulting in a sink for the fluvial

J.M. Edmond is in the Department of Earth and Planetary Sciences at the Massachusetts Institute of Technology.

Lonsdale⁹ had found evidence in 1977 of active springs in the area of high heat flow in the Guyamas Basin in the northern part of the Gulf of California (see *News and Views* 289, 9; 1981). This portion of the ridge axis is buried under several hundred metres of sediments principally supplied by the Colorado River. Subsequent Deep Sea Drilling Holes (leg 64) showed the presence of basaltic sills in the sediment column, some of them still hot¹⁰. In the summer of 1980 a Deep Tow survey yielded photographs of active hot springs; sulphide minerals were recovered by dredging and the plastic liner of a core barrel melted at a depth of several metres in the sediment. Recovered cores were lithified and contained very high concentrations of light hydrocarbons produced by thermal cracking of buried planktonic debris¹¹. Venting of very hot waters is clearly occurring despite the sealing effect of the thick sediment blanket.

Galápagos-type hydrothermal mounds were discovered in Deep Sea Drilling Program holes (leg 60) on the flanks of the back-arc spreading centre in the Marianas

Basin¹². An extinct massive sulphide deposit was found at a ridge-transform intersection on the Galápagos ridge during dives by *Alvin* at the beginning of 1980¹³.

The joint US-French programme on the East Pacific Rise also met with success. In the summer of 1980 an expedition on the Jean Charcot using the 'Seabeam' high-resolution bathymetric mapping array and the RAIE towed camera system found active sulphide deposition at several locations. Numerous samples were obtained by dredging¹⁴. Major diving campaigns are planned in the Eastern Pacific in 1981 using both US and French submersibles. □

1. Corliss, J. B. *et al.* *Science* 203, 1073 (1979).
2. Williams, D. L. *et al.* *J. geophys. Res.* 84, 7467 (1979).
3. Edmond, J. M. *et al.* *Earth planet. Sci. Lett.* 46, 1 (1979).
4. Edmond, J. M. *et al.* *Earth planet. Sci. Lett.* 46, 19 (1979).
5. Jenkins, W. J., Edmond, J. M. & Corliss, J. B. *Nature* 272, 156 (1978).
6. CYAMEX Scientific Team *Nature* 277, 523 (1979).
7. RISE Project Group *Science* 207, 1421 (1980).
8. Edmond, J. M. *EOS* 61, 992 (1980).
9. Lonsdale, P. F. *et al.* *Earth planet. Sci. Lett.* 49, 8 (1980).
10. Einsele, G. *et al.* *Nature* 283, 441 (1980).
11. Lonsdale, P. F. *EOS* 61, 995 (1980).
12. Hegarty, K. A. *et al.* *EOS* 61, 364 (1980).
13. Malahoff, A. *et al.* *EOS* 61, 208 (1980).
14. Boulegue, J. *et al.* *EOS* 61, 992 (1980).

Monitoring Mount St Helens

from M. McCormick

IN November, just six months after the 18th May 1980 cataclysmic eruption of Mount St Helens, Washington, a multidisciplinary group of researchers brought their various viewpoints and data to Washington DC*. Among the problems they discussed were: Was the eruption going to affect climate or local weather? What are the important measures describing a volcanic eruption — height of the cloud, sulphur content, lava flow, or other factors? How did Mt St Helens compare with past volcanic eruptions? The uniqueness of the meeting came from the variety of disciplines and research areas represented. Volcanologists, geologists, biologists, meteorologists, atmospheric chemists and physicists, climate modellers, remote sensor specialists and geochemists were all working together.

When Mt St Helens erupted, there existed, for the first time, sufficiently sophisticated technology to measure the global movement and enhancement of stratospheric aerosol mass loadings. This, coupled with a mature modelling capability, a number of programs specifically set up to measure volcanic effluents and stratospheric aerosol radiative characteristics, and a readily accessible volcano, all helped to make this

volcano and its eruptions the best documented in history.

The NASA orbiting satellites SAM II (Stratospheric Aerosol Measurement) and SAGE (Stratospheric Aerosol and Gas Experiment), developed specifically for the global measurement of stratospheric aerosols, were important new sources of information. Their international ground truth programs helped drive the development of a number of new lidar systems which were in operation at the time of the eruption. These lidar systems, specifically designed for stratospheric aerosol measurements, are in Japan, West Germany, Italy, France, and the US. They were complemented by programs that used aircraft platforms and a number of *in situ* sensors — for example, the NASA ACE (Aerosol Climatic Effects) program which began the year before and which provided another set of instruments poised and ready to measure stratospheric aerosol parameters important to climate, just as the volcano began its rumblings. The NASA-sponsored RAVE (Research on Atmospheric Volcanic Emissions) is an even newer program whose acceptance was in final negotiations when Mt St Helens erupted on 18 May and which is designed to visit volcanoes close up, and characterize their source function, atmospheric chemistry, and potential impact on climate.

As Mt St Helens is in the US a continuous meteorological network was available so that detailed trajectory analyses could be performed. The Geosynchronous Operational Environmental Satellite-West was also in position to generate electronically pictures of the early heavy plumes and their movements.

Material in the stratosphere has the potential to alter the radiative transfer of solar and terrestrial radiation. This in turn can cause warming or cooling of the surface and atmosphere. The radiative characteristics of the material, such as index of refraction and size distribution, and the underlying albedo, determine the net effect. Model calculations, based on past eruptions and stratospheric aerosol characteristics, have typically shown a warming in the stratosphere and cooling at the surface.

During periods of no major volcanic activity, stratospheric background aerosols are thought to be composed primarily of sulphuric acid and water in an approximate 75/25 mix. Violent volcanic eruptions impart enough energy to the ejecta to penetrate the local tropopause. Once in the stratosphere, the gaseous and ash-like components of the eruption cloud can have a long lifetime as in this region of the atmosphere the rate of cleansing is slow. Thus, material can disperse and spread globally and alter climate. The heavier ash-like or non-volatile particles can fall out faster than the gaseous material (primarily various sulphur gases). The sulphur gases are converted to aerosol particles which are thought to be identical to the background $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$ aerosol. The new aerosols are distributed in a size range about 0.1–0.6 μm which is effective in scattering visible sunlight. The distribution of large particles closer to 1.0 μm is not well known at present but is also important in understanding the radiative transfer through the layers. Based on past data, the lifetime (the time for its mass to reduce to 1/e of its value) of new volcanically-produced stratospheric aerosols at 18-km altitude is about one year. How long it takes for the sulphur gases to convert to $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$ droplets is uncertain, but at the workshop evidence was presented that it takes place quickly, within a month. There is evidence that some conversion even occurred in the erupting plume or in the magma itself. The amount of sulphur gas that is put into the stratosphere becomes crucial to the determination of an eventual effect on climate. Early aircraft measurements showed the aerosol in the vertical column on 18 May to be primarily composed of silicates of larger size below 16 km with sulphuric acid-clad silicates above. Balloon-borne measurements of the material over Wyoming showed a

*The Mt St Helens Symposium was held from November 18 to 21, 1980, chaired by J.P. Friend of Drexel University, Philadelphia and was followed by a workshop chaired by R.E. Newell of MIT.

M. McCormick is Head of the Aerosol Measurements Research Bureau at the Langley Research Center, Virginia.

predominance of non-volatile stratospheric particles early after the 18 May eruption. After about a month, the particles were volatile and similar to background aerosols in their characteristics. Lidar data at two colours and with a measurement at crossed polarizations, were consistent with this conclusion.

Material from the eruption column over Mt St Helens was moved by a system of winds whose direction and speed vary with altitude. A large percentage of the upper atmospheric material was at jet stream altitudes between about 8 km and 12 km, and moved at average speeds of about 20 m s^{-1} circling the globe in an eastward path in as little as 16 days. Higher material moved much more slowly and in sometimes very complex trajectories. For example, material between 15 and 16 km moved eastward, crossing the eastern US coast on 23 May, and reaching Europe no later than 5th June. The material above 20 km was above the wind reversal, and moved westward at an average speed of about 6 to 8 m s^{-1} . Local radar and other remote sensors showed this material reached at least 23 km and travelled around the globe in about 56–70 days. Ground-based lidars tracked this upper altitude material as it crossed westward from country to country.

Most of the radiative models that determine temperature changes for an assumed set of aerosol radiative characteristics describe the temperature change that is caused by a certain mass of new aerosol. The determination of mass, therefore, is very important in estimating the temperature change. Both SAGE and lidar data have been converted through optical modelling to aerosol mass and used to determine the mass added to the background stratosphere. SAGE data yielded a value of 0.25 to 0.5×10^6 metric tons and lidar 0.4 to 0.6×10^{10} metric tons. These calculations are for July when the material began to disperse, affecting the entire Northern Hemisphere. SAGE's 'global eye' showed the material to be most highly concentrated at latitudes greater than 50° N . A normal back-ground hemisphere contains about 0.25×10^6 tonnes of aerosol and the addition due to the eruption therefore represents an approximate 100 per cent increase. Radiative models predict that this increase, although significant, would only cool the near surface temperature by hundredths of a degree and warm the stratosphere near the peak aerosol concentration at 18 km by a degree. The exact values depend, of course, on the specific aerosol characteristics and are in the process of being determined but will not be great enough to have a significant effect on climate. In comparison, the eruption of Agung on Bali was thought to have introduced about 30×10^6 metric tons into the stratosphere in 1963, causing a reported 0.3° C cooling at the surface and about a 4 to 8° C warming in the stratosphere. □

Will the real Grenville Orogeny please stand up

from Keith Bell

THE PROTEROZOIC remains more of a mystery than any other period of the Earth's history and is currently presenting several tantalizing problems, particularly regarding large-scale crustal movement and the formation of orogenic belts. It is generally considered to be a period (2,500 Myr to about 600 Myr ago) of crustal rifting, shearing and dyke emplacement with major transcurrent lineaments, marking large-scale lateral displacements^{1,2}. Because of this, and the scarcity of ophiolites and greenstone belts, some geologists believe that the tectonic regime during the Proterozoic differed from anything that existed either before or since.

The Grenville Province of Canada is one of the most perplexing of the Proterozoic orogenic belts and the numerous models suggested for its origin reflect not only a great deal of speculation but also a marked lack of detailed field mapping. Although the Grenville belt shows features apparently quite different, at present exposure levels, from Phanerozoic continental margin mountain belts (for example, the virtual absence of linear belts of basic igneous rocks and the presence of late alkaline complexes and carbonatites), it still remains unclear whether the pervasive 1,000 Myr Grenville 'event' developed intracratonically or as a result of a collision event involving plate motion. Geological evidence tends to favour the latter^{3–6}.

In this issue of *Nature* (p.129), Baer draws attention to the significance or lack thereof of the 1,000 Myr date for the Grenville 'event' and proposes not one but two orogenies, an earlier one at about 1,100 Myr and a later one at about 950 Myr. Reflected in this paper are many of the problems inherent in the interpretation of geochronological data from polymetamorphosed terranes and the difficulties involved in placing constraints on the so-called Grenville Orogeny.

Although generalization about discordant age patterns in orogenic belts is difficult, most can be attributed to differences either in the temperatures at which the daughter nuclides become 'frozen' into the system during cooling, or in the abilities of different minerals to become open systems to the various daughter nuclides during thermal overprinting. Because of the relative ease of argon migration out of most systems, it is difficult to distinguish slow cooling from overprinting solely on the basis of K–Ar dates. The geological interpretation of the data often depends on the degree of discordancy. According to Baer, the 150 Myr

Keith Bell is Professor of Geology at Carleton University, Ottawa, Canada.

gap between the 'younger' K–Ar and the 'older' U–Pb and Rb–Sr dates from the Grenville Province is too long to be explained by slow cooling and his preference lies with a metamorphic overprinting at about 950 Myr. Others, however, take the opposite view^{7,8}. Although Baer's interpretation of the length of time and the comparison of K–Ar dates with those from other systems might be open to criticism, at least two events can be proposed solely on the basis of the U–Pb and Rb–Sr dates.

Baer's conclusions pose two major questions — what do we mean by the term Grenville Orogeny, and how many events have affected the Grenville Province? The original subdivision of the Canadian Shield into areas affected by distinct orogenies (Kenoran $\sim 2,500$ Myr, Hudsonian $\sim 1,740$ Myr, Elsonian $\sim 1,400$ Myr, Grenvillian $\sim 1,000$ Myr), using the mean value of many K–Ar mineral (mainly mica) dates⁹, successfully outlined large regions of the Shield that had experienced similar terminal thermal histories. But what do these ages really mean? If the term 'orogeny' is used in the sense of epeirogenesis, then the argon dates can be interpreted as dating the time of uplift and cooling, an event which is sometimes, but not always, substantiated by sedimentary sequences of the same or slightly younger age. However, there is no good reason to assume that crustal movements are synchronous both across and along the length of an orogenic belt. Fault-bounded crustal segments ('zones', 'domains') that are now recognized in many orogenic belts of different ages (for example, the Appalachian and Hudsonian orogenies)^{10,11} appear to have had their own distinct thermotectonic histories which may differ from one another by some tens of millions of years. The time of cooling to relatively low blocking temperatures, such as those for argon in micas, can thus vary considerably within a single orogenic belt and because of this it is no longer acceptable to define an orogeny solely on the blanketing together of large numbers of K–Ar dates. The term 'orogeny' in its present usage, in any case, means something more than simple uplift and encompasses the more geologically significant processes of folding, metamorphism and intrusive activity.

The resolution of events in polymetamorphosed terranes requires both isotopic data from two or more radiometric systems and good geological control based on sound field mapping. Questions such as 'Are the dated bodies pre-, syn- or post-tectonic?' and 'How many foliations can be observed in the field?' will have direct

bearing on the interpretation of the isotopic data. The real problem is relating isotopic to geological events, although this may be impossible either because the geology is not known in sufficient detail or because the isotopic event has no visible expression in the rocks themselves.

Unravelling the extremely complex history of the Grenville Province is difficult because the geology of the belt is poorly understood. Some parts have never been mapped at other than the 1:1,000,000 scale although those mapped in detail have yielded rich rewards. A well defined stratigraphy in the relatively low-grade Hastings Basin region demonstrated a major unconformity that can be correlated over 2,000 km², separating a younger psammitic, pelitic and carbonate sequence (the Flinton Group) from an older volcanic, sedimentary and igneous terrane. The known field relationships coupled with the isotopic data pinpoint the deposition of the Flinton succession to between 1,050 and 1,080 (± 25) Myr ago. Furthermore, three deformations (including the one that produced the prominent north-east-trending structures) and an accompanying regional metamorphism all post date 1,050 Myr (ref. 12). Although this represents a fairly small chapter in the history of the Grenville Province, it does emphasize the need for assigning both structural and stratigraphical positions to the various sedimentary and metavolcanic units within the belt and relating them to existing or new isotopic data.

The history of the Grenville Province is much more complicated than that discussed by Baer whose compilation deals with only the last two hundred million years or so. Both geological and isotopic evidence demonstrate its continued tectonic overprinting extending back into the Archaean, including the effects of both the Hudsonian and Elsonian orogenies¹³.

Recent whole-rock Rb-Sr isochrons indicate a granulite facies event at about 1,430 Myr and perhaps two magmatic events at 1,250 and 1,120 Myr from a relatively restricted area of the belt^{14,15}. The questions now remain of how many orogenies have affected the Province and which, if any, should be designated the 'Grenville Orogeny'. This sort of problem is not restricted to the Grenville Province alone. Magmatic and metamorphic events from the Saskatchewan segment of the Churchill Province at 2,430, 1,880, 1,740 and 1,560 Myr have recently been recorded by U-Pb zircon and whole-rock Rb-Sr studies^{16,17}. What then is the age of the Hudsonian Orogeny?

Despite a case having been made for using the term 'Grenville' only in a stratigraphical sense¹⁸, most workers would probably like to use it more generally or even as a specific name for one of the orogenies. If one takes the latter approach, the widespread event at about 1,050 $\pm$ 50 Myr that produced or was superimposed on major north-east-trending structures is a likely candidate. It has all the hallmarks of a well defined orogeny, has been substantiated by more than one radiometric method and is fairly widespread throughout the central part of the Province. An appealing alternative proposes the term 'Grenvillian orogenic cycle', analogous to the Appalachian orogenic cycle that includes the Taconian,

Acadian and Alleghanian orogenies. Already two orogenies, the Elzevirian and the Ottawan have been proposed.

The period from 1,150 to 1,000 Myr ago was undoubtedly a geologically active one, not only in the Grenville Province, but elsewhere in the Canadian Shield. Widespread basic magmatism (Logan sills and Keweenaw lavas in the Lake Superior region, basaltic flows and diabase dykes in the Northwest Territories) and deep-seated alkalic and carbonatitic activity in northern Ontario¹⁹⁻²² suggest a time of large-scale rifting and faulting. Perhaps a closer look at Proterozoic sequences and intrusive bodies outside the Grenville Province may help to pin down more accurately the age of the main Grenville event. Whether such an approach will hold further clues about the origin of the Grenville Province and its deformation or whether such ages are merely coincidental will have to be substantiated by further work.

So the Grenville thorn still remains embedded in the side of the geologist interested in the Proterozoic. Baer's paper, however, certainly forces a re-evaluation of the significance of the $\sim 1,000$ Myr figure so often quoted for the Grenville Orogeny. Whether the two proposed events at $\sim 1,100$ Myr and ~ 950 Myr are geologically significant will only become clear as more detailed field mapping emerges. □



100 years ago

LAST Thursday the Hackney Microscopical and Natural History Society held their annual *soirée*, always a very successful event. Many other similar London societies were represented at the meeting.

An elaborate report upon the opening up of two of the pyramids at the boundary of the Libyan Desert near Sakkara is now published by Prof. Brugsch. The learned professor estimates the matter to be of the most important and valuable kind. At the close of 1880 the entrances to the sepulchral chambers of the three pyramids were laid bare. The ceilings were taken off, and only the two sides, all covered with hieroglyphics, rose from the *débris*. The hieroglyphics point to the reign of Pharaoh Apappus.

An important experiment in electric lighting is about to be made in the City. Hitherto the electric light has been used, as on the Thames Embankment and Waterloo Bridge, in conjunction with gas; but in the City the thoroughfares selected are to be lighted by electricity alone, which will be continued all

night. The experiment is to be continued for a year, at an outlay of about 8000*l*.

It is known that the young horns of the *Cervus maral* (Severtzoff), when they are filled with blood and not yet ossified, are very much prized by the Chinese, who purchase them at the Siberian frontier, paying as much as six to twenty pounds the pair. A very active chase of the maral has therefore always been carried on in Siberia, and since it became rather rare, the Cossacks in the neighbourhood of Kiakhta have domesticated this stag. Now we learn from a communication by M. Polakoff that its domestication has greatly extended in Western Siberia, so that there are herds of seventy head; but the horns of the domesticated deer, as might be expected, have lost a good many of their original qualities.

A stenographic piano has been experimented on by the daughter of the inventor, in the French Chamber of Deputies, the Senate, and to the Municipal Council of Paris, with great success. The system consists of a combination of signs through which every sound is represented. The reproduction is as rapid as speaking, and the same operator can continue the work for hours. The signs used in this system being printed by machinery, the reading is immediate, and can be made by other people than the operator. The State stenographers propose to be trained in the use of the instrument. It is an affair of a few months of practice.

from *Nature* 23, 3 and 10 March 1880.

1. Sutton, J. & Watson, J.V. *Nature* 247, 433 (1974).
2. Onstott, T.C. & Hargraves, R.B. *Nature* 289, 131 (1981).
3. Dewey, J.F. & Burke, K.C.A. *J. Geol.* 81, 683 (1973).
4. Brown, R.L., Chappell, J.F., Moore, J.M. & Thompson, P.H. *Geosci. Can.* 2, 141 (1975).
5. Baer, A.J. *Phil. Trans. R. Soc. A* 280, 499 (1976).
6. Young, G.M. *Earth Sci. Rev.* 16, 277 (1980).
7. Berger, G.W., York, D. & Dunlop, D.J. *Nature* 277, 46 (1979).
8. Dallmeyer, R.D. & Sutter, J.F. *J. geophys. Res.* 85, 3177 (1980).
9. Stockwell, C.H. *Geol. Surv. Pap. Can.* 64-17, 1 (1965).
10. Williams, H., Kennedy, M.J. & Neale, E.R.W. *Geol. Ass. Can. Pap.* 11, 181 (1972).
11. Lewry, J.F. & Sibbald, T.I.I. *Can. J. earth Sci.* 14, 1453 (1977).
12. Moore, J.M. & Thompson, P.H. *Can. J. earth Sci.* 17, 1685 (1980).
13. Wynne-Edwards, H.R. *Geol. Ass. Can. Pap.* 11, 263 (1972).
14. Davidson, A., Britton, J.M., Bell, K. & Blenkinsop, J. *Geol. Surv. Pap. Can.* 79-1B, 153 (1979).
15. Bell, K. & Blenkinsop, J. *Geol. Surv. Pap. Can.* 80-1C, 152 (1980).
16. Ray, G.E. & Wanless, R.K. *Can. J. earth Sci.* 17, 333 (1980).
17. Macdonald, R. & Broughton, P. (compilers) *Geological Map of Saskatchewan*, Prov. Edn. 1:1,000,000 (Saskatchewan Geological Survey, 1980).
18. Gilluly, J. *Am. J. Sci.* 264, 97 (1966).
19. Gittins, J., MacIntyre, R.M. & York, D. *Can. J. earth Sci.* 4, 651 (1967).
20. Douglas, R.J.W. (ed.) *Geol. Surv. Can. Econ. Geol. Rep.* 1 (1970).
21. Hanson, G.N. *Can. J. earth Sci.* 12, 821 (1975).
22. Bell, K., Blenkinsop, J. & Watkinson, D.H. *Abstr. GAC/MAC a. Mtg Quebec* (1979).

REVIEW ARTICLES

Textons, the elements of texture perception, and their interactions

Bela Julesz

Bell Laboratories, Murray Hill, New Jersey 07974, USA

Research with texture pairs having identical second-order statistics has revealed that the pre-attentive texture discrimination system cannot globally process third- and higher-order statistics, and that discrimination is the result of a few local conspicuous features, called textons. It seems that only the first-order statistics of these textons have perceptual significance, and the relative phase between textons cannot be perceived without detailed scrutiny by focal attention.

THE study of pre-attentive (also called effortless or instantaneous) texture discrimination can serve as a model system with which to distinguish the role of local texture element detection from global (statistical) computation in visual perception. Indeed, the question of whether there are basic elements in perception out of which higher percepts are built, or whether these higher percepts are indivisible, as the Gestaltist psychologists claimed, is still debated. Without using the sophisticated techniques described in this article, it is not obvious, even in the case of pre-attentive texture discrimination, whether local differences between the texture elements directly contribute to discrimination or whether these differences are sensed in a global way only through differences in the statistics of the texture.

A few examples will clarify this problem. For instance, the texture pair in Fig. 1a effortlessly separates into two distinct textures with a sharp boundary between them. This is true even when the texture pair is presented for 160 ms (or less), outside foveal gaze (to prevent scrutiny of the textures element-by-element by scanning eye movements), and when erasure is used afterwards (to prevent scrutiny of the texture pair's after-image by shifts of focal attention). One 'explanation' of this pre-attentive texture discrimination is that the local texture elements in the two textures which have different sizes are processed by local detectors selectively tuned to different sizes; texture discrimination would be directly conveyed by these detectors. Alternatively, one could describe the difference between the two textures globally, by the difference in their first-order statistics, and assume that this statistical difference underlies

pre-attentive texture discrimination. The first-order statistic (or probability distribution) is simply the probability that randomly thrown dots will land on a certain colour (for example, black) of the texture. This first-order statistic is clearly very different for the two textures in Fig. 1a.

Similarly, the effortless texture discrimination in Fig. 1b can be easily explained by the local orientation difference between the elements that constitute the two textures. But again it is possible to describe the difference between the texture pair globally. Here, the first-order statistics of the two textures are identical, but their second-order statistics differ greatly. The second-order statistic (or joint probability distribution) is the probability of the events that the vertices (end points) of randomly thrown 2gons (dipoles or needles) of all possible lengths and orientations will fall on certain colours of the texture (for example, both on black). [The n gon statistic (or n th order joint probability distribution) of an image can be obtained by randomly throwing n gons of all possible shapes on the image and observing the probabilities that their n vertices fall on certain colour combinations.] Obviously, the second-order (or dipole) statistics are very different for the two textures in Fig. 1b. Finally, Fig. 1c is an example in which the textures are defined globally by keeping their second-order statistics different but their first-order statistics identical, yet the resulting discriminable textures also exhibit locally perceivable differences. Here the left half-image is generated by a two-dimensional Poisson process, while the right-half field is generated by the same process, but with the restriction that no two dots can fall closer to each other than some distance ϵ . A local basis for texture

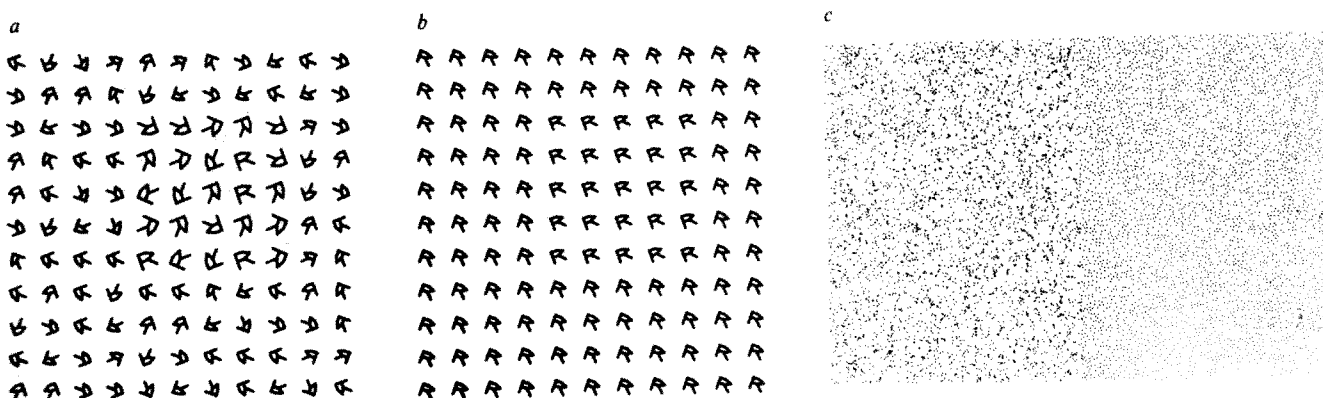


Fig. 1 Pre-attentively distinguishable texture pairs based either on global statistics or local conspicuous elements, or both: a, with different first-order statistics and difference in element size; b, with different second-order statistics and different element orientation; c, with different second-order statistics that result in short neighbouring dots (dipoles) of critical size that can occur only in one half-field. (From B.J. *et al.*¹⁰.)

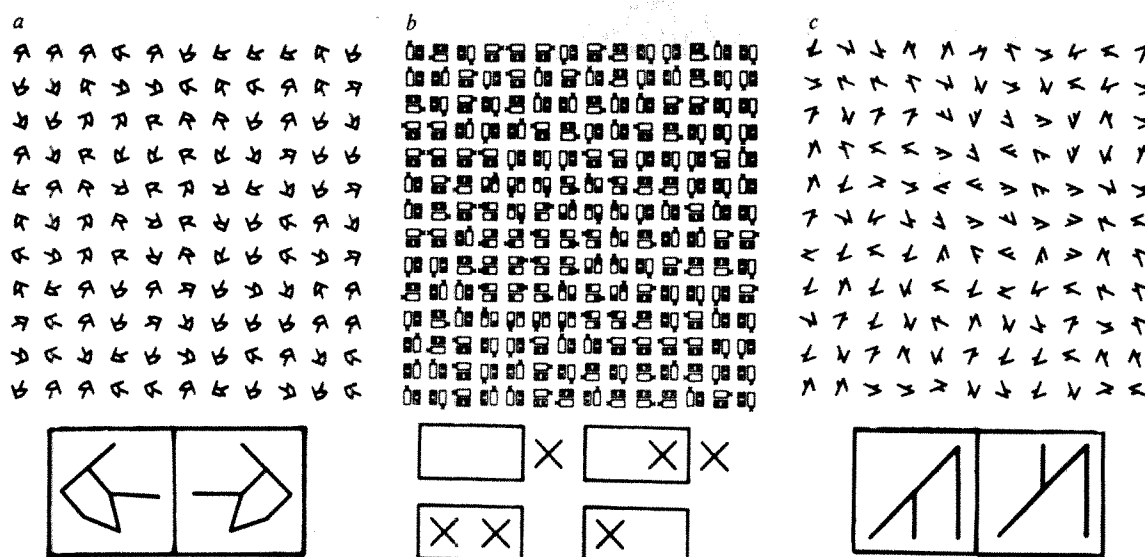


Fig. 2 Pre-attentively indistinguishable texture pairs with identical second-order but different third- and higher-order statistics: *a*, composed of randomly thrown similar micropatterns and their mirror images, respectively (from B.M. *et al.*¹⁰); *b*, composed of dual micropatterns, as described in Fig. 3*b* (from B.J.¹⁶); *c*, composed of dual micropatterns as described in Fig. 3*b* (from B.J.¹⁶).

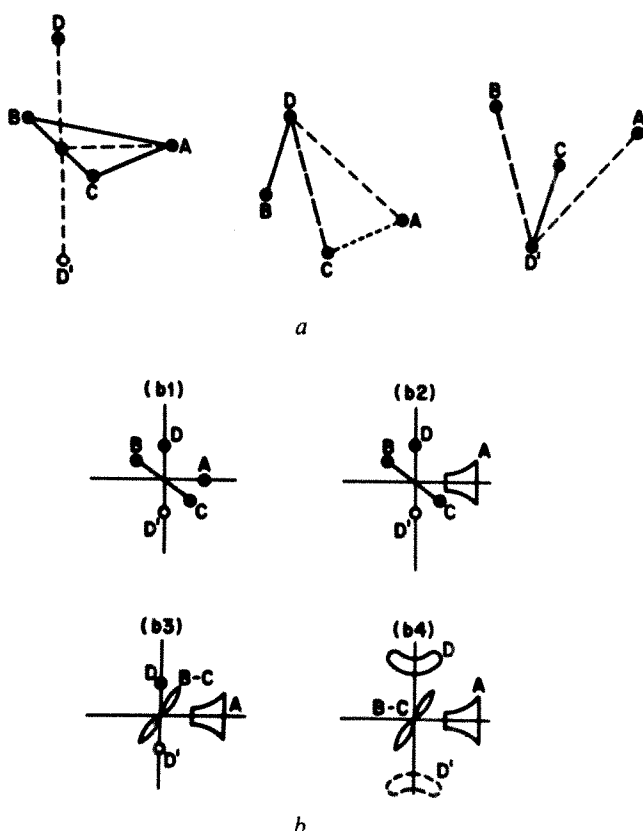


Fig. 3 The four-disk method and its generalization to generate isometric dual micropatterns that when randomly thrown, respectively, yield texture pairs with identical second-order statistics. *a* Depicts the four-disk method of B.J. *et al.*¹⁰. One micropattern is composed of the four identical disks A, B, C and D, such that A is centred on the x-axis, B and C are centric symmetric to the origin, while D is centered on the y-axis. The dual micropattern contains D' in place of D, that is symmetric with respect to the origin of D. The dual micropatterns are isometric, but there exist triangles (ABD) that cannot fall on the dual micropattern (after B.J. *et al.*¹⁰). *b*, Generalization of the four-disk method of four steps B1, B2, B3 and B4, where disk A can be any object with bilateral symmetry across the x-axis. BC disks can be replaced by any centric symmetric object, and D and D' can be any objects with bilateral symmetry across the y-axis, and centric symmetric to each other (after Caelli *et al.*¹³).

discrimination exists if one considers feature detectors with receptive fields (of ϵ or smaller extent) as are found in the visual system of vertebrates¹⁻⁴. Texture segregation could be explained by the frequent stimulation of these detectors in the left field, whereas no such detectors are activated in the right field.

These examples indicate that effortlessly discriminable textures with different second- or first-order statistics—which include most, if not all, natural textures—also seem to differ locally. So whether the pre-attentive visual system merely detects the occurrence of certain local features that are present in one texture and absent in another, or analyses some global statistical parameter that distinguishes these textures is still undecided. At this point, some readers might favour the theory of feature detectors, influenced in their choice by the dramatic neurophysiological discoveries of such detectors in the visual systems of frog, cat and monkey by such workers as Kuffler¹, Barlow², Lettvin *et al.*³ and Hubel and Wiesel⁴. However, Barlow, who pioneered the notion of a receptive field in neurophysiology², was unable to find psychophysical evidence for elongated receptive fields in human texture perception. He recently studied the detection of one rectangular random-dot array surrounded by another having different first-order statistics⁵ and found, to his surprise, that performance of detection did not change with the aspect ratio of the target rectangle, despite the fact that most receptive fields in the monkey cortex are highly elongated⁶. Indeed, it took 16 years to find evidence that local feature detectors participate in texture perception.

The iso-second-order texture paradigm

Figure 1*a*, *b* and *c* demonstrate that the pre-attentive visual system might process differences in second-order statistics, or just detect conspicuous local differences, or perhaps performs both operations. In 1962, to test the limit of statistical computational capabilities of the pre-attentive visual system, I posed the question: would it be possible to create texture pairs with identical second-order statistics, but different third- and higher-order statistics; and could such iso-second-order textures still be discriminated? The mathematicians Rosenblatt and Slepian solved the first part of this question for Markov processes⁸, and in probing the second part of the question, I observed that iso-second-order Markov textures of densely packed stochastic dot arrays were usually indistinguishable without scrutiny⁷. Recently, a new class of two-dimensional, non-Markov stochastic dot textures has been created by Pratt *et al.*⁹, and these were also indistinguishable when their second-order statistics

agreed. From these observations one could conjecture that "the pre-attentive textural system cannot globally compute third- or higher-order statistics". This will be called the 'modified Julesz conjecture'. (The original Julesz conjecture, which hypothesized that "iso-second-order textures are indistinguishable", has been disproved and will be discussed below.)

The main problem with these indistinguishable iso-second-order stochastic dot textures is that they are also locally indistinguishable. It would be more convincing to test the modified Julesz conjecture using iso-second-order texture pairs whose elements are conspicuously different yet as textures are indistinguishable. Recently I and my co-workers¹⁰⁻¹⁴ have succeeded in generating several classes of iso-second-order texture pairs that are globally (that is, as textures) indistinguishable but locally (that is, their elements) can be discriminated. Figure 2a, b and c demonstrate three such texture pairs.

The texture pair in Fig. 2a is composed of identical non-overlapping elements ('R') and their mirror-images in all possible random orientations. Such texture pairs, as I and my co-workers¹⁰ have shown, have identical second- but different third- and higher-order statistics. (Indeed, when one of these textures is printed on a glass sheet, the other side shows the mirror-image texture, and obviously the dot and dipole statistics must agree, but the triangle statistics differ.) The texture pair in Fig. 2a is sufficiently indistinguishable that it is unlikely, without prior prompting, that one would suspect that this texture pair is not a single coherent texture. One has to scrutinize the textures element by element and perform a 'mental rotation' in the Shepard and Metzler sense¹⁵. Indeed, even when the pair of elements is presented in isolation, the two elements cannot be discriminated in a 160-ms brief flash followed by erasure if they are oriented far from parallel.

Figure 2b depicts an even more interesting case. This iso-second-order texture pair and many others were generated by the generalization of the four-disk method described by myself and my co-workers¹² as developed by Caelli *et al.*¹³. Figure 3 legend describes how the four-disk method and its generalization work. One can prove the geometry that the micropattern composed of the four disks ABCD, and its partner composed of the disks ABCD' are isometric: to any two points (dipole)

selected in one micropattern there exists a corresponding dipole in the dual micropattern. Although these dual dipoles usually have different orientations, the textures formed by the random throwing of the micropatterns in all orientations result in identical dipole statistics. The generalization of the four-disk method is as follows. One can easily verify that disk A can be chosen to be any shape with bilateral symmetry to the x-axis. B and C disks can become any shape with centric symmetry through the origin, while D and D' can be replaced by any shape with bilateral symmetry across the y-axis and equally distant from the origin. The dual micropatterns composed of these three forms again are isometric. However, there are triangles whose three vertices will fall on one micropattern, but not on its dual micropattern. Thus, the texture pair in Fig. 2b has the same second-order, but different third-order statistics. Here the elements of the two indistinguishable textures are pre-attentively indistinguishable although in this case with focal attention they can be easily told apart. After all, one would expect that simple cortical units whose elongated receptive fields match the rectangles in these texture elements should differentially fire for zero, one or two X's being contained in these rectangles. That no textural discrimination is obtained suggests two possibilities: (1) differences in third- or higher-order statistics are not processed; and (2) the output of local feature analysers is linearly averaged. The latter statement refers to the observation that areas covered by rectangles containing only single X's when linearly averaged have the same contribution as rectangles containing zero or two X's when these occur with 0.5 probability and one assumes that the simple cortical units are themselves linear.

Figure 2c is perhaps the most instructive example. It is also generated by the technique described in Fig. 3b. Here the indistinguishable iso-second-order texture pair is composed of dual elements, respectively, that in isolation are pre-attentively discriminable. Thus, when the dual elements are briefly flashed (followed by erasure) and the two elements placed 6 deg arc apart (3 deg arc from the fixation point) they can be discriminated, regardless of their orientations. The lack of texture discrimination in Fig. 2c is rather unexpected. If one were to assume that Kuffler units and elongated blob detectors of the Hubel and Wiesel type are involved in perception—as attested

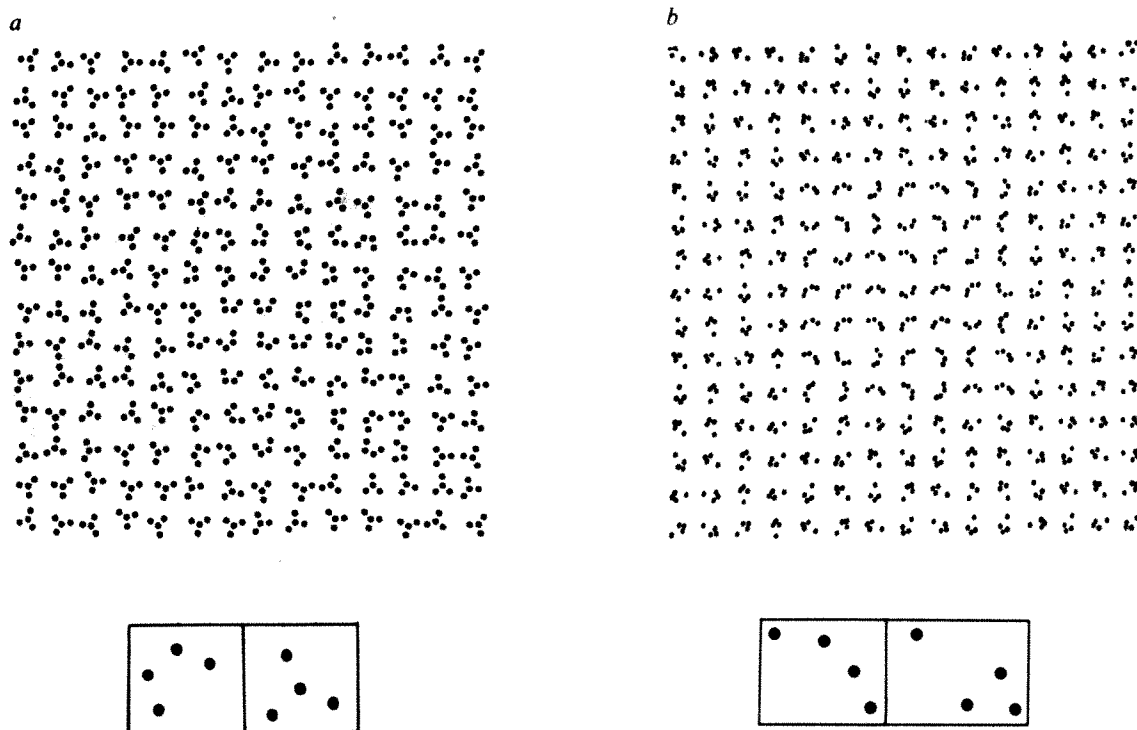


Fig. 4 Iso-second-order texture pair generated by the four-disk method: a, just below threshold of pre-attentive texture discrimination (from B.J. *et al.*¹⁰); b, well above threshold. Pre-attentive texture discrimination is based on local conspicuous features of quasi-collinearity of disks. (From Caelli and Julesz¹².)

to by the fact that the 'open' and 'compact' dual elements can be discriminated—it is rather surprising that pools of such detectors (differentially tuned to the larger open and smaller compact elements) are not utilized in pre-attentive texture discrimination. This problem will be resolved in the following sections.

It had been shown that iso-second-order textures exist that are indistinguishable globally even when locally their elements are either attentively, or even pre-attentively, distinguishable. Because the second-order statistics determine the autocorrelation function, and the Fourier transform of the autocorrelation function is the power spectrum, these iso-second-order textures have identical autocorrelation functions and identical power spectra. Thus, the pre-attentive visual system ignores the phase (position) spectrum that is so conspicuous when the elements are viewed by focal attention. This implies the operation of two visual systems, as has been suggested by many workers in the fields of attention and cognition. The attentive system uses focal attention scanned serially. The pre-attentive system uses distributed attention that is mediated by a parallel process. The insensitivity of pre-attentive texture perception to the position of texture elements directly implies a parallel system.

The role of local elements in iso-second-order texture discrimination

The existence of many indistinguishable iso-second-order textures has established that the pre-attentive visual system is unable to process statistical information beyond the second order. The next question is whether iso-second-order texture pairs do exist that yield pre-attentive texture discrimination. If they do, such texture discrimination must be the result of local features. As we have seen that many seemingly conspicuous features do not yield texture discrimination in the iso-second-order statistical constraint, those local features that might yield texture discrimination should be very special ones. They should be regarded as the basic elements of pre-attentive perception, and I have named such hypothetical elements 'textons'¹⁶.

Indeed, in 1978 several new iso-second-order texture classes were discovered that yield strong discrimination based on very conspicuous local features. The first was discovered by Caelli

and myself¹² using the four-disk method given in Fig. 3a. In Fig. 4a a case is shown (using the dual micropatterns shown in Fig. 3a) that is just below the threshold of pre-attentive texture discrimination, although the open and closed texture elements in isolation are pre-attentively discriminable. However, it was possible to open up the open elements further so that the four disk micropatterns became quasi collinear (while keeping the dual elements still 'closed'). As shown in Fig. 4b, presenting the quasi-collinear features in only one texture region resulted in strong pre-attentive texture discrimination. This local feature might be visually extracted by local detectors tuned to elongated blobs (of given width, orientations and aspect ratios).

Several other iso-second-order textures have since been found that yield pre-attentive texture discrimination^{12-14,16-18}. In each case, pre-attentively discriminable local conspicuous features such as 'closure', 'corner', 'connectivity' and 'granularity' were prominent. Figure 5a, b and c show a few typical iso-second-order texture pairs that yield pre-attentive texture discrimination.

Figure 5a and b were obtained by the generalized four-disk method of Fig. 3b, and discrimination seems to be due to the conspicuous local features of closure and connectivity, respectively. Figure 5c is an iso-third-order (hence also iso-second-order) texture pair created by a new stochastic process of Julesz, Gilbert and Victor¹⁴ which shows that, contrary to what might be believed, perceived granularity need not be determined by the power spectrum, nor even by the second- and third-order statistics but can be a fourth-order property! Here again, as in the case of quasi-collinear disks, it seems that elongated blobs of different widths, length and orientation, and perhaps of different first-order statistics in small, local neighbourhoods are the local conspicuous features yielding texture discrimination.

The role of terminators in texture discrimination

Only the quasi-collinear structures (line segments) and elongated blobs in general have so far been identified as textons. Are other conspicuous local features such as closure, corner and connectivity (Fig. 5a, b) also textons? Fortunately, I have recently found that this is not so¹⁶. In one experiment, dual

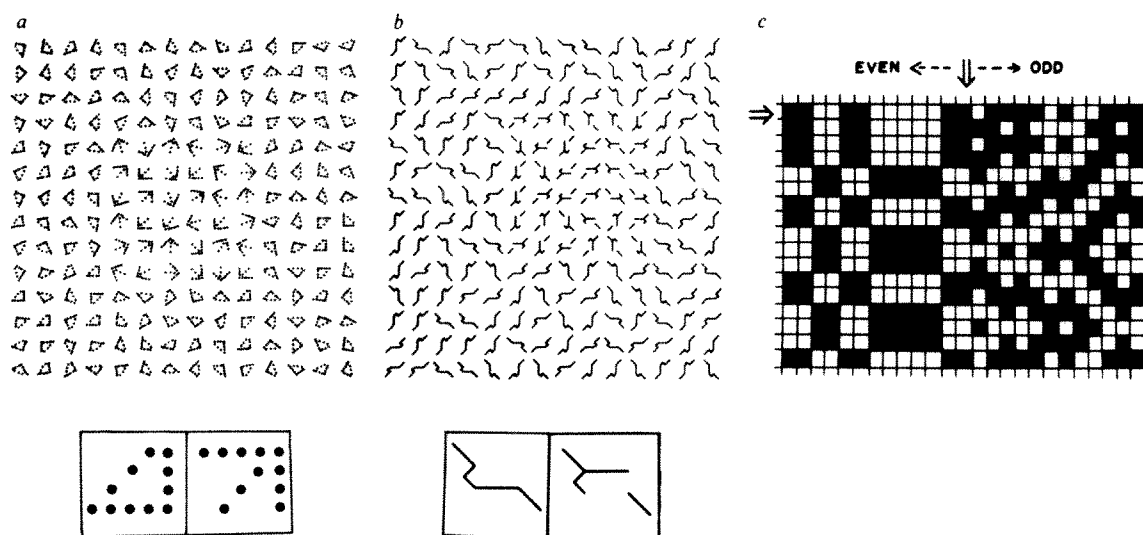


Fig. 5 Discriminable iso-second-order texture pairs based on local conspicuous features of: a, 'closure' (from Caelli *et al.*¹³); b, 'connectivity' (from B.J.¹⁶; a and b were generated by the method described in Fig. 3b); c, 'granularity'. This texture pair has identical third-order (hence identical second-order) statistics (from B.J. *et al.*¹⁷). The small squares in the first row and middle column are selected black and white at random while each 2×2 square left of the middle column contains even numbers of black squares and right of the middle column contains odd numbers of black squares. However, two basic local features (textons) suffice to explain these differences. In c textons are elongated blobs of different sizes, while for a and b it is the difference in the number of end-of-lines (terminators) of the dual elements that yields texture discrimination.

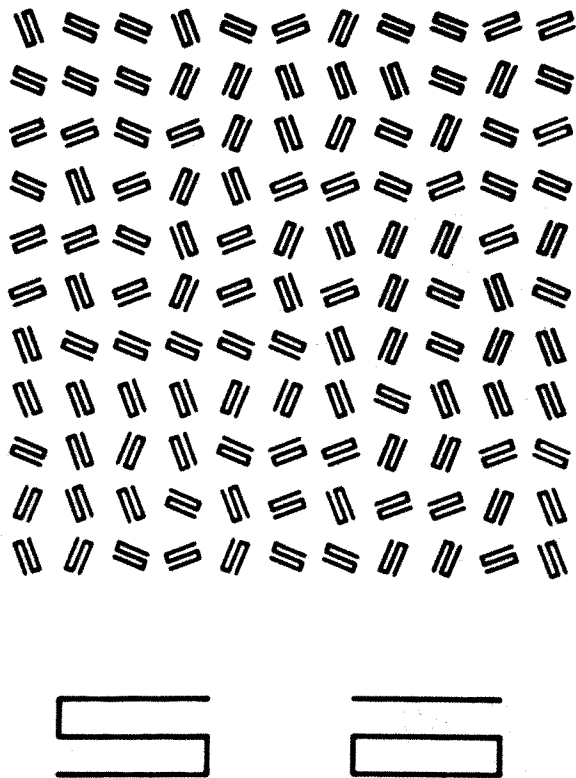


Fig. 6 Demonstration that if the number of terminators (end-of-lines) agrees, then in pre-attentive texture discrimination their exact positions are ignored. This texture pair is indistinguishable in spite of the fact that the open/closed and connected/unconnected features are juxtaposed in the dual elements; however, line segments and the number of terminators agree.

texture elements were selected in which open/closed and connected/unconnected features were simultaneously juxtaposed (insert in Fig. 6). When these formed texture pairs they could not be pre-attentively discriminated from the surround (Fig. 6). One might suspect that the texture pair in Fig. 6 is indistinguishable (similar to the indistinguishable pair in Fig. 2c) because, in addition to having the same line segments, the number of the end points (terminators) of these line segments is also the same in the dual texture elements. Furthermore, the fact that the texture pair in Fig. 6, composed of 'S' and '10'-shaped elements, is indistinguishable, also shows that the exact position of the terminators is not perceived pre-attentively.

Indeed, if one inspects the discriminable iso-second-order texture pairs in Fig. 5a and b, the basis of discrimination can be seen to be the difference in terminator number. The pre-attentive visual system, evidently, cannot determine the location of terminators, but can count their numbers (or density) or their first-order statistics. Thus, the putative textons of corner, closure and connectivity can actually be simply described by differences in terminator numbers.

Therefore, besides colour, there are two additional texton classes: elongated blobs (of given orientation, width and aspect ratio) and their terminators. Pre-attentive texture discrimination is based either on the difference in the textons or the difference in the first-order statistics (or some parameter defined by the first-order statistic, such as density or standard deviation) of the textons.

Thus, the original Julesz conjecture has been refuted: there are iso-second-order textures that can be discriminated. However, the modified Julesz conjecture seems to hold: the pre-attentive textural system cannot compute statistical parameters higher than second order. Therefore, discrimination

in iso-second-order textures involves differences in local conspicuous features, called textons.

Interaction between textons

One can even strengthen the modified Julesz conjecture by raising the possibility that the pre-attentive textural system globally cannot even compute second-order statistical parameters, but can evaluate only the first-order statistics of textons. If this were the case, then for the pre-attentive system the interactions between textons would be simple, based on some first-order statistical parameters.

An experiment, depicted by Fig. 7 (and published here for the first time) suggests such a possibility. In Fig. 7a the same micropattern comprised of a 4×4 dot array of randomly selected 8 white and 8 black dots is repeated many times horizontally and vertically with an 8-dot period (leaving 4-dot-wide blank gaps between the micropatterns). This periodically covered area is flanked on two sides by an area of densely packed random dots of black and white. Figure 7b is derived from Fig. 7a by filling the blank gaps between the periodic micropatterns with densely packed black and white random dots. The texture pair in Fig. 7b (composed of the area with periodic micropatterns and the neighbouring random area) have identical first-order statistics but very different second-order statistics (and very different autocorrelations). Nevertheless, the texture pair appears indistinguishable. On the other hand, if the 4×4 dot micropattern contains a texton (consisting of 8 white and 8 black dots forming stripes) as shown in Fig. 7c and d (with and without blank gaps between the periodic micropatterns, respectively), Fig. 7d yields strong texture discrimination, based on the dense occurrence of textons that are practically absent in the other texture. Indeed, the largest difference in the autocorrelations between the periodic and aperiodic textures occurs at the 8-dot periods, and this difference is the same for the indistinguishable texture pair in Fig. 7b and the strongly distinguishable one in Fig. 7d. These examples clearly indicate that texture discrimination is not the result of differences in second-order statistics, but is based on density changes in textons. Caelli and myself¹⁹, who studied discrimination of texture pairs composed of dot pairs (dipoles) having different orientational ranges, could predict perceptual performance based on the first-order statistic of dipole orientation alone.

It is interesting that if texture elements and their dual partners differ in their textons, even a single element can be pre-attentively perceived as being different amidst many iso-second-order dual partners (Fig. 8a).

Recently, Burt and I became interested in the interactions between textons, or interactions between the first-order statistics of different textons²⁰. We were particularly interested in cooperative interactions, and to this end we tachistoscopically presented two or more target micropatterns embedded in iso-second-order dual partners, either dispersed or grouped, as shown in Fig. 8a and b. We found that if the critical (unshared) texton belonged to the target micropattern and not the surround, there was no cooperativity. The dispersed and grouped conditions were equally detected, and probability of detection increased with the number of target micropatterns (that is, with the number of the critical textons). Also, probability of detection was greater when the critical texton belonged to the target micropattern instead of the surround. In another texture paradigm, Beck²¹ noted that the presence of features (textons) in the target is more perceptible than their absence.

If the critical texton(s) are absent from the target micropatterns (and thus belong to the surround), then the grouped condition is better perceived than the dispersed. However, this 'cooperativity' is of a rather simple kind. If we assume pools (layers) of similar textons corresponding to various retinal positions, then the presence of a critical texton in a target micropattern will cause activity in the critical pool (and increased activity with many targets, according to their number) regardless of whether these textons are dispersed or grouped. However, the absence of critical textons in the target micropatterns means

that these textons belong to the micropatterns in the surround, and densely stimulate the corresponding texton layer. To be certain that a texton-free area is not a chance fluctuation, this area must be significantly larger than the mean distance between the micropatterns of the surround. Obviously, grouped target micropatterns assure a much larger texton-free area than do dispersed ones.

Finally, we also observed that with long practice, involving several hundred trials, some indistinguishable iso-second-order texture pairs (such as Fig. 2c) could become pre-attentively discriminable. This might seem contradictory to the recent cognitive studies of Treisman²², who conjectured that conjunctions of features (she used colours and shapes) cannot be pre-attentively perceived, even with practice, but require serial, element-by-element search. The research reported here agrees with her findings. Only disjunctive texton groups (that is, differences in textons) can be perceived pre-attentively (that is, as a parallel process, regardless of element number). The reason the conjunctions of textons that form the dual elements in Fig. 2c can be learned might be that these open and closed elements can have different stimulatory effects on elongated blob detectors,

which are themselves textons. Their detection escaped attention before the task was mastered.

Textons and Marr's 'primal sketch'

It was only after 18 years that I and my co-workers discovered how the simple, complex and hypercomplex cortical feature detectors of the neurophysiologists would correspond to the texton detectors of elongated blobs and their terminators. Thus, it took that long to overcome the initial scepticism, and show that the feature detectors of the neurophysiologist could be utilized in pre-attentive perception. Marr²³, within the framework of artificial intelligence (machine vision), proposed his 'primal sketch' model based on local intensity changes across edges, elongated blobs (lines) and end-of-lines, and their first-order statistics. In essence, the input is spatially filtered and convoluted with masks of given shapes, widths and orientations, similar to the 'Mexican hat function' shape receptive fields of simple cells of Hubel and Wiesel⁶. Marr also introduced inhibition between such simple units (particularly of perpendicular

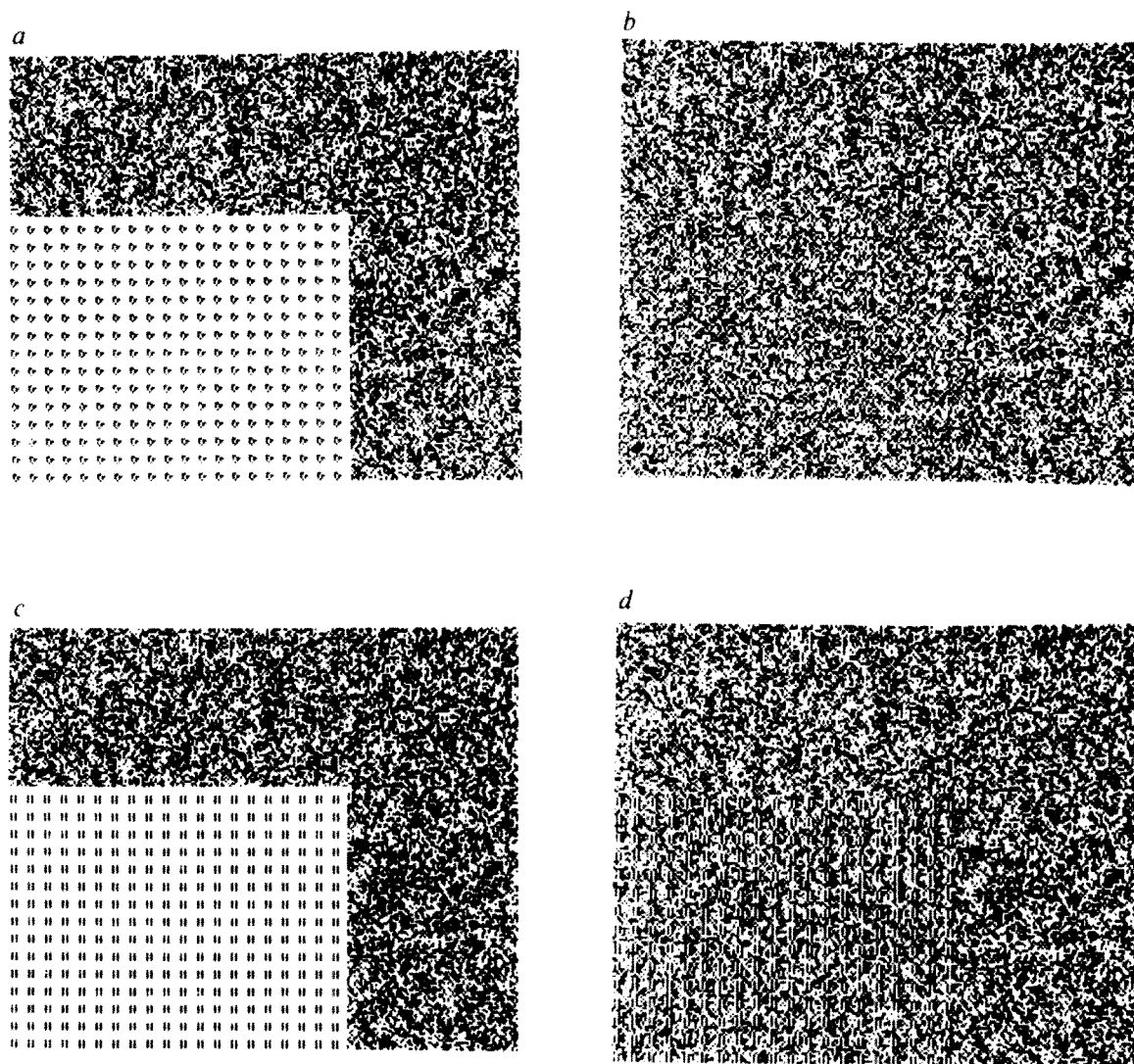


Fig. 7 Demonstration that the pre-attentive textural system cannot process globally differences in second-order statistics, and only first-order statistics of textons yield texture discrimination. *a*, Periodically repeated 4×4 dot micropattern that does not contain textons, flanked by randomly dotted areas. *b*, Same as *a*, except that the blank gaps between the periodic micropatterns are filled with random dots. The second-order statistics are very different between the area composed of periodic micropatterns and its random neighbour, and yet they appear indistinguishable. *c*, Same as *a*, except that the 4×4 dot micropattern contains textons of periodic stripes. *d*, Same as *c*, except that the noise insertion described in *b* is carried out. Now strong texture discrimination is experienced based on the dense occurrence of textons that are missing in the neighbouring area.

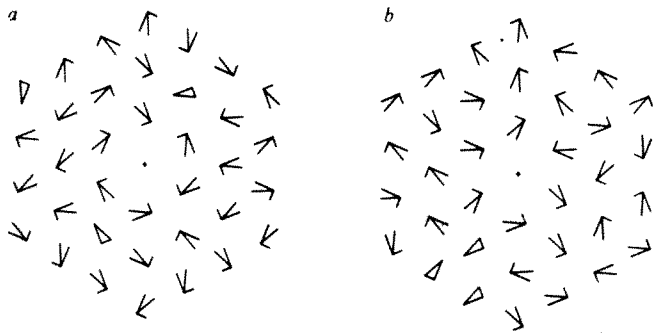


Fig. 8 Demonstration that texton differences can be pre-attentively detected even if the target micropatterns are dispersed within many dual elements as in *a*; in *b* a few target elements are grouped. With centre fixation and briefly flashed (using erasure), the dispersed and grouped case yields identical detection when the unshared textons belong to the target.

units) giving rise to terminators, trying to emulate all the known feature detectors of the neurophysiologists, and devised algorithms that would emulate pre-attentive vision. Thus, Marr was the first to stress the role of terminators in texture perception. However, he went a step further and introduced 'place tokens', higher-order collections of blobs (lines) based on several grouping operations that do not depend on semantic content. Space does not permit a review of this important model, but let me quote Marr (ref. 23, p. 510): 'Julesz (1975, p. 43) mentioned in an aside the possibility that texture vision may rest on 'first-order statistics of various simple feature extractors' but this idea requires the concept of the primal sketch and of recursively applied grouping before it can be brought into fruition.' Although Marr's primal sketch model goes beyond texture perception, into pre-attentive vision and beyond, neither he nor neurophysiologists or researchers in artificial intelligence could have guessed that in texture discrimination only the number of terminators has perceptual significance and that their position is ignored. Similarly, without careful psychological research, one cannot guess how real and virtual dipoles (between two place tokens) affect perception. For example, the primal sketch model regards them as equal, which led to incorrect predictions in texture discrimination as shown by Schatz²⁴. Also, the finding that in texture perception only three or more quasi-collinear place tokens act as a real line shows the importance of perceptual experiments¹⁶. It remains to be seen whether, besides the three texton classes, some other textons will be found in pre-attentive texture discrimination—the more complex place-tokens of Marr. On the other hand, anyone who wants to go beyond texture perception should carefully study Marr's discerning ideas.

Obviously, artificial intelligence, neurophysiology and psychology can aid each other, and the fact that the search for textons led to findings that agree with ideas and discoveries in artificial intelligence and neurophysiology is most gratifying.

Conclusion

It was shown that contrary to the Gestaltists' belief, a highly global percept, texture, can be decomposed into elementary

units, the texton classes of colours, elongated blobs of specific widths, orientations and aspect ratios, and the terminators of these elongated blobs.

One might ask whether sinusoidal gratings, so popular in vision research, could be used in texture perception. After all, the detection of compound sinusoidal gratings does not depend on phase (if the ratio of the gratings is 3:1 or higher), as shown by Graham and Nachmias²⁵. This phase independence at threshold of detection, however, should not be confused with the phase independence of suprathreshold discrimination often found in iso-second-order textures. While textures with identical second-order statistics have identical power spectra, the converse is not true. One cannot profitably study textures having only identical power spectra, because changes in phase between sinusoidal grating components dramatically change the first-order statistics, and thus give rise to texture discrimination for trivial reasons. Furthermore, global Fourier analysis cannot reveal the occurrence of local textons, particularly terminators²⁶ (also implied by the demonstrations of Fig. 7 and ref. 28).

Thus, through the iso-second-order texture paradigm we can restrict the global computational powers of pre-attentive perception and identify local conspicuous features, the textons. Whatever the final number of these texton classes will be, it is evident that for texture discrimination the Gestaltists' view has not been corroborated: the global percept of texture was found to be decomposable into more basic elements. Only if pre-attentively discriminable texture pairs were to be discovered, with its elements (or local aggregates of these elements) found to be pre-attentively indistinguishable, would my modified conjecture be refuted.

This article was restricted mainly to research based on the iso-second-order texture paradigm. The reader interested in more conventional texture research, particularly in the more complex problems of textural grouping and recognition, and in practical problems such as that of automating the search for malignancy in clinical tissues, might find additional references in ref. 27.

In conclusion, let me draw an analogy to the phase-transitions of physics, in which symmetry (invariance) laws and the breaking of these symmetries prevail. Textons in pre-attentive perception obey such a symmetry law. They seem, within limits, to be invariant under positional and scaling transformations. These symmetries of textons cannot be broken by pre-attentive and peripheral vision, as the 'coupling' that locks textons in their proper positions is missing (as if the 'attentional temperature' in these systems were too high). Only by inspecting areas where the densities of textons differ, thus switching to focal attention, can our visual system break the texton symmetries. Only within a small disk of focal attention are the textons coupled (as if this disk were 'cooled' below a critical temperature), permitting the formation of higher percepts (thus enabling us to distinguish, say, an R from its mirror image). This analogy mainly serves to pinpoint the major implication of our findings, that most of the single cortical feature analysers (outside foveal representation) might not be connected directly to each other (but interact only in aggregate). This relative phase insensitivity might also account for the apparent poor spatial resolution of our peripheral vision, which, however, is quite acute for textons, particularly end-of lines.

- Kuffler, S. W. *J. Neurophysiol.* **16**, 37–68 (1953).
- Barlow, H. B. *J. Physiol., Lond.* **119**, 69–88 (1953).
- Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. & Pitts, W. H. *Proc. Inst. Radio Engrs* **47**, 1940–1951 (1959).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **148**, 574–591 (1959).
- Barlow, H. B. *Vision Res.* **18**, 637–650 (1978).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215–243 (1968).
- Julesz, B. *IRE Trans. Inf. Theory* **IT-8**, 84–92 (1962).
- Rosenblatt, M. & Slepian, D. *J. Soc. ind. appl. Math.* **10**, 537–549 (1962).
- Pratt, W. K., Faugeras, O. D. & Gagalowicz, A. *IEEE Trans. Syst. Management Cybernet.* **SMC-8** (11), 796–804 (1978).
- Julesz, B., Frisch, H. L., Gilbert, E. N. & Shepp, L. A. *Perception* **2**, 391–405 (1973).
- Julesz, B. *Scient. Am.* **212**, 34–43 (1975).
- Caelli, T. M. & Julesz, B. *Biol. Cybernet.* **28**, 167–175 (1978).
- Caelli, T. M., Julesz, B. & Gilbert, E. N. *Biol. Cybernet.* **29**, 201–214 (1978).
- Julesz, B. in *Perceptual Organization* (eds Kubovy, M. & Pomerantz, J. R.) (Erlbaum Associates, Hillsdale, New Jersey, in the press).
- Shepard, R. N. & Metzler, J. *Science* **171**, 701–703 (1971).
- Julesz, B. *Phil. Trans. R. Soc. B* **290**, 83–94 (1980); also reprinted in *The Psychology of Vision* (eds Longuet-Higgins, H. C. & Sutherland, N. S.) 83–94 (R. Soc. Lond., 1980).
- Julesz, B., Gilbert, E. N. & Victor, J. D. *Biol. Cybernet.* **31**, 137–140 (1978).
- Victor, J. D. & Brodie, S. E., *Biol. Cybernet.* **31**, 231–234 (1978).
- Caelli, T. M. & Julesz, B. *J. opt. Soc. Am.* **69**, 675–678 (1979).
- Julesz, B. & Burt, P. *Bull. Psychonomic Soc.*, 242 (1979).
- Beck, J. J. *Perception Motor Skills* **36**, 1331–1341 (1973).
- Treisman, A. in *Levels of Processing in Human Memory* (eds Craik, F. I. M. & Cermak, L. S.) 301–330 (Erlbaum Associates, Hillsdale, New Jersey, 1978).
- Marr, D. *Phil. Trans. R. Soc. B* **275**, 483–524 (1976); in *The Psychology of Vision* (eds Longuet-Higgins, H. C. & Sutherland, N. S.) 199–218 (R. Soc. Lond., 1980).
- Schatz, B. *Computer Sci. Dept Rep.* **152**, (Carnegie-Mellon University, Pittsburgh, 1978).
- Graham, N. & Nachmias, J. *Vision Res.* **11**, 251–259 (1971).
- Julesz, B. & Caelli, T. M. *Perception* **8**, 69–73 (1979).
- Julesz, B. & Schumer, R. A. *A. Rev. Psychol.* **32**, 575–627 (1981).
- Julesz, B. *Biol. Cybernet.* (in the press).

Estimate of risk from environmental exposure to radon-222 and its decay products

R. D. Evans*, J. H. Harley†, W. Jacobi‡, A. S. McLean§, W. A. Mills||
& C. G. Stewart¶

* Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Environmental Measurements Laboratory, US Department of Energy, New York, New York 10014, USA

‡ Institut für Strahlenschutz, Gesellschaft für Strahlen- und Umweltforschung mbH, München, FRG

§ National Radiological Protection Board, Harwell, Didcot, Oxfordshire OX11 0RQ, UK

|| Criteria and Standards Division, US Environmental Protection Agency, Washington DC 20460, USA

¶ Chalk River Nuclear Laboratories, Chalk River, Ontario, Canada KOJ 1P0

Consideration of the epidemiological evidence on radiogenic lung cancer in uranium miners and of the incidence of the disease generally leads to an upper estimate for the lifetime risk of 10^{-4} cases per working level month for members of the general population.

IN radiological protection, the emphasis is usually on limiting exposure to artificial radiation. Thus, there is a well-established approach to the protection of staff in hospitals from X rays and to the protection of members of the public from radioactive consumer goods or from effluent from nuclear power plants. How one should deal with exposure to natural radiation is less clear.

Human beings have always been exposed to natural radiation both from within and outside the body. The view adopted by the International Commission on Radiological Protection (ICRP) is that exposure to normal levels of natural radiation cannot usually be subject to control, but it is acknowledged that there may be unusual circumstances in which control is desirable¹. The Commission has in mind circumstances in which there is a considerable increase in exposure as a result of human practices or choices of environment.

One of the principal ways in which members of the public receive natural irradiation is by breathing the radon-222 decay products in air. Radon gas is emitted by soil, rocks and building materials, all of which contain the parent radium to some extent. Radon can also be present in water, from which it is readily desorbed. Out of doors, the gas is dispersed in the atmosphere, but it concentrates indoors to a degree dictated primarily by the ventilation rate of the dwelling and the rate of radon input. The solid short-lived decay products of radon grow in rapidly² and, if attached to an aerosol rather than to walls or furniture, are inhaled and irradiate the lung tissues. Lung cancer may result.

Even in normal circumstances, human exposure to radon decay products is substantial compared with other components of the overall exposure to natural radiation. In abnormal circumstances, radon exposure may be dominant and perhaps also unacceptably high. Such abnormal circumstances stem from the use of building materials with relatively high radium content, for example, or the use of domestic water with exceptionally high radon content.

The basis of dose limitation in radiological protection is risk limitation. For example, the dose equivalent limit for workers is set so that the average risk likely to be incurred is comparable with the average prevailing for other workers in industries regarded as having high standards of safety. For members of the public, the level of acceptability of risk is probably an order of magnitude lower. It is important, therefore, to have some information on risk factors for environmental exposure to radon decay products if a system of control is to be considered.

Direct risk information does not exist for environmental levels of radon decay products. It is therefore necessary either to extrapolate from data for occupational exposure or to infer the

risk from the incidence of lung cancer in the wider population. Both approaches are followed here. We have taken care neither to exaggerate the risk nor to overextend the evidence.

Uranium miner data

The only occupational data that can be used for obtaining a risk estimate for lung cancer caused by radon decay products are those from the exposure of underground uranium miners. Although the BEIR report³ does give an estimate of the risk of lung cancer from whole-body external irradiation, this is insufficient for the present purpose. Of the four sets of data on miners, those from Sweden and Canada are considered to be less useful, and we have therefore concentrated on the data from Czechoslovakia and the United States. Information on exposure is given in terms of working level month (WLM), the unit commonly employed in mining. One WLM is defined as exposure for 1 working month of 170 h to a concentration of one working level (WL), where one WL is any combination of short-lived decay products of radon-222 per litre of air that will result in the emission of 1.3×10^5 MeV of α energy during complete decay^{2,4,5}.

Several factors cause the data on miners to be more difficult to use than might at first appear. The average concentrations throughout US mines in the 1940s and 1950s were high (7–15 WL) and mean exposures accumulated over the 1950s and 1960s were about 800 WLM (ref. 6); in the last decade, however, levels were lower by one to two orders of magnitude^{7,8}. Most of the early information on exposure was obtained by the operators, frequently in areas where ventilation problems were known to exist: on the other hand, there was sometimes a tendency for higher exposures to be under-reported. All the measurements on aerosol characteristics, such as particle size and unattached fraction, have been made under modern mining conditions. These conditions include improved ventilation, the presence of diesel smoke (which probably overwhelms other particulates) and the virtual disappearance of the one-man mine. The evaluation of risk is also complicated by the presence of other carcinogens in mine air as well as by the miners' smoking history and their previous experience of hard-rock mining.

The estimates of lifetime risk that emerge from the data available on miners range from one estimate of 21–54 deaths from lung cancer per 10^6 WLM of collective exposure to one of 1,000 or so. Stewart and Simpson (unpublished) and Myers and

Stewart⁹ have evaluated the American and Czech data using various statistical techniques. Their work indicates that the incidence of lung cancer can be accounted for by a linear relationship with exposure, allowing a constant factor for non-radiogenic lung cancers; their estimate is 21–54 lung cancers per 10^6 WLM from the US data¹⁰, but the Czech data imply a risk about three times as great¹¹. The discrepancy is not readily explained. Stewart and Simpson also found a real or apparent threshold in some of the analyses, which suggests that the estimate of risk for low-level exposure may even include zero as a lower bound. (In such cases, Stewart and Simpson used the slopes of the response curves to derive risk factors.) Jacobi¹² also noted the discrepancy between the US and Czech data and proposed a range of 100–500 lung cancers per 10^6 WLM, which is virtually identical with an earlier estimate by UNSCEAR¹³; in neither case is a threshold effect posited. The highest estimate of risk of about 1,000 lung cancers per 10^6 WLM comes from Archer's recent proposal¹⁴.

Whereas it is not possible completely to rule out any of these estimates of the lifetime risk from occupational exposure, our objective is to estimate risk to the general population from exposure to radon decay products in the general environment. In using the data for uranium miners to derive the lifetime probability of fatal lung cancer for members of the public, one must note that there are significant environmental and physiological differences between the two situations. These differences result from the possible contribution to lung cancer induction in mining environments of other dusts and gases, from differences in breathing rates, from differences in equilibrium ratios of the radon decay products, from high as opposed to low concentrations, from differences in smoking habits and from population distributions that differ in age and sex. Although some of these factors may tend in opposite directions, on balance, they suggest that members of the general public may be at a lower risk per unit exposure than miners. For example, breathing rates for a man resting indoors could be two to three times lower than for a man mining, with the concomitant decrease in exposure counteracted to some degree by an increase in the probability of depositing radon decay products in the lung.

We judge from the epidemiological evidence that the most defensible upper bound of the lifetime risk to the general population is 100 deaths from lung cancer for a collective exposure of 10^6 WLM, or 10^{-4} per WLM. This particular factor, although severely rounded to avoid the impression of unwarranted accuracy, can be supported by other indirect epidemiological evidence, which is detailed below, and by an analysis (W. Jacobi, personal communication) of the dosimetry of radon decay products in the manner of Jacobi and Eisfeld¹⁵.

Lung cancer incidence

Most calculations of the exposure of the general population were initially made on the basis of outdoor radon levels. Indoor levels were known to be higher, but it is only recently that an appreciable number of measurements has become available. UNSCEAR¹³ summarized the literature appearing before 1976, and more detailed measurements have since been published^{16–19}. Information on living habits in the Northern Hemisphere indicates that more than 90% of the day is spent indoors, so that indoor exposure is used here to represent total exposure.

The activity concentration of radon indoors in the Northern Hemisphere lies, on the average, in the range of 0.5–1 pCi per l. If a value of 0.8 pCi l^{-1} is assumed, some measurements in the United States¹⁶ indicate that this can be converted to a decay product concentration of $\sim 0.004 \text{ WL}$. For 168 h per week of indoor occupancy, the exposure would be 0.2 WLM in 1 yr or 12 WLM in 60 yr. The latter period might be shortened somewhat, but this would not change our general conclusion.

Given this mean lifetime exposure of 12 WLM and the risk factor of 10^{-4} per WLM given above, a maximum lifetime risk of 0.12% is obtained by extrapolating linearly and by neglecting latency and accumulation interval. This is a relatively small fraction of the present lifetime risk of lung cancer (about 4% in the United States), which is largely attributable to cigarette smoking.

Making the reasonable assumption that exposure of the general population to radon decay products has been constant over several decades, one can compare the estimate of 0.12% with the incidence of lung cancer before the effects of smoking became pronounced. There are relatively good data²⁰ on incidence in the United States in 1930 which yield a lifetime risk of 0.1%, roughly 0.15% for males and 0.05% for females. Since not all lung cancers at that time would have been caused by radon, this indirect evidence supports the view that a value of 10^{-4} per WLM is a reasonable upper bound of risk to members of the public.

Quite independently, Cliff and others^{17,18} have studied the relationship between the observed exposure to radon decay products and the observed incidence of lung cancer in British women. Taking an average exposure rate of 0.15 WLM per yr, and assuming a risk factor of 2×10^{-4} per WLM (ref. 6) as well as a 20-yr latency period, they predict a higher incidence of lung cancer than is observed from all causes in women below 40 years of age. This supports the evidence from the United States, and indeed an upper bound of 10^{-4} per WLM for the lifetime risk would not be incompatible with the British epidemiological evidence.

The overall evidence therefore points to an upper estimate for the lifetime risk of 10^{-4} cases per WLM for members of the general population.

Risk estimates compared

It is of interest to compare the risk of lung cancer from exposure to radon decay products with that from external penetrating radiation. On the basis of a linear non-threshold model, and purely for protection purposes, ICRP estimates that the risk factor for external irradiation of the lung alone is 2×10^{-3} per Sv or 2×10^{-5} per rem, where Sv and rem are the new and old units of dose equivalent, respectively¹. Thus, the lifetime risk from an exposure to radon decay products of 1 WLM, namely 10^{-4} , would not exceed that from an external irradiation dose equivalent of 50 mSv (or 5 rem) to the lung alone.

It is difficult, however, to imagine circumstances in the general environment in which the lung alone would be exposed to external radiation: whole-body irradiation is much more likely. The ICRP takes the view that most, if not all, forms of fatal malignant disease and hereditary damage can be induced by exposure to ionizing radiation and that, when all the organs of the body are uniformly exposed, cancer of the lung in particular will account for about 10% of all harmful effects. Thus, the lifetime risk associated with an exposure to radon decay products of 1 WLM would not exceed the overall risk associated with a dose equivalent of about 5 mSv (or 0.5 rem) to the whole body from external radiation.

The annual dose equivalent limit recommended by the ICRP for individual members of the public is in fact 5 mSv.

This review has its origins in an international workshop on radiation protection principles for naturally occurring radionuclides held at Arlington, Virginia, in May 1978 by the Nuclear Energy Agency of the OECD. The authors, who were members of a group considering radon problems, have had a number of subsequent exchanges and discussions on the estimation of risk from this source. Thus the present paper should be attributed to them, even though they had the advantage of the initial group deliberations which included M. Kinoshita. The authors also acknowledge the assistance they received from M. C. O'Riordan.

1. *Recommendations of the International Commission on Radiological Protection*, Publ. 26 (Pergamon, Oxford, 1977).
2. Evans, R. D. *Hlth Phys.* **17**, 229–252 (1969).
3. BEIR Report *The Effects on Populations of Exposure to Low Levels of Ionizing Radiation*, 372–397 (National Academy of Sciences, National Research Council, Washington DC, 1980).
4. Kusnetz, H. L. *Am. ind. Hyg. Ass. Q.* **17**, 85–88 (1956).
5. *Federal Register* **40**, 50704–50705 (1975).
6. Jacobi, W. *Proc. NEA Specialist Meet. Personal Dosimetry and Area Monitoring Suitable for Radon and Daughter Products*, Elliot Lake, Canada, 33–48 (Nuclear Energy Agency, OECD, Paris, 1977).
7. Richardson, H. P. *Proc. NEA Specialist Meet. Personal Dosimetry and Area Monitoring Suitable for Radon and Daughter Products*, Paris, 45–51 (Nuclear Energy Agency, OECD, Paris, 1979).
8. Swent, L. W. *Testimony at Hearings on the Impact on Older Workers of Occupational Health Hazards*, (US Senate Committee on Aging, Grants, New Mexico, 1979).
9. Myers, D. K. & Stewart, C. G. *AECL-5970* (Chalk River Nuclear Laboratories, 1979).
10. Lundin, F. E., Wagoner, J. K. & Archer, V. E. *NIOSH and NIEHS Joint Monogr.* **1** (NTIS, Springfield, 1971).
11. Ševc, J., Kunz, E. & Pláček, V. *Hlth Phys.* **30**, 433–437 (1976).
12. Jacobi, W. *Proceedings of NEA Specialist Meeting on Personal Dosimetry and Area Monitoring Suitable for Radon and Daughter Products*, Paris, 223–230 (Nuclear Energy Agency, OECD, Paris, 1979).
13. *United Nations Scientific Committee on the Effects of Atomic Radiation, 1977 Report to the General Assembly, Annexes G & B* (United Nations, New York, 1977).
14. Archer, V. E. *Trans. Am. nucl. Soc.* **33**, 145–146 (1979).
15. Jacobi, W. & Eisefeld, K. *GSF-Report S-626* (Institut für Strahlenschutz, München, 1980).
16. Breslin A. J. & George, A. C. *Trans. Am. nucl. Soc.* **33**, 150–152 (1979).
17. Cliff, K. D. *Radiol. Prot. Bull.* **22**, 18–23 (1978).
18. Cliff, K. D., Davies, B. L. & Reissland, J. A. *Nature* **279**, 12 (1979).
19. Rundo, J., Markun, F. & Plondke, N. J. *Hlth Phys.* **36**, 729–730 (1979).
20. *Ca-A Cancer Journal for Clinicians*, 18–19 (1976).

ARTICLES

Structure, strength, and polarization changes in radio source SS433

R. M. Hjellming

National Radio Astronomy Observatory, Socorro, New Mexico 87801, USA

K. J. Johnston

E. O. Hulburt Center for Space Research, Naval Research Laboratory, Washington DC 20375, USA

VLA observations of SS433 at 1,465, 4,885, and 15,035 MHz between 8 September 1979 and 20 June 1980 show structure with size scales between 0.1 and 5 arc s which changes significantly on time scales of 1–2 weeks. Jets with 8–20% linear polarization are ejected on both sides of the central radio source which is coincident with the optical image, along position angles between 80° and 120° with a mean position angle for ejected material of 100°. Features in linear polarized intensity maps can be identified in four successive epochs spanning 196 days, and the resulting proper motions for identifiable clouds are 0.0088 arc s per day. If the feature velocities are equal to that of the optical emission lines, the distance to SS433 is 5.1 kpc. The position angles of identifiable features correlate well with the known precessing optical jets if one extrapolates back to the time of ejection using the observed transverse speeds. The equipartition magnetic field strengths of 0.001–0.01 G indicate synchrotron lifetimes of 10^9 – 10^{11} s; therefore synchrotron losses are not a dominant factor in the evolution of the extended radio emission. The evolution of radio jets indicates that the relativistic electrons and magnetic fields seen as synchrotron emission may be generated by flows along the rotation axis of the accretion disk of the SS433 star system.

DISCUSSIONS^{1–4} of the peculiarities in the optical emission lines of SS433 together with various theories^{5–7} have established that SS433 has oppositely directed jets of optically emitting material moving with speeds of 0.26c. Variations in the apparent Doppler shifts of emission lines indicate an oscillation in the direction of the jets which can be interpreted as precession of an accretion disk responsible for the weak X-ray emission⁸ and the high velocity flows perpendicular to the accretion disk which

appear as the optically observed jets. The detection of radio emission from SS433 (refs 9, 10) has led us to question whether SS433 has radio jets analogous to those in extragalactic radio sources or double radio sources such as those associated with the strong X-ray source Sco X-1 (ref. 11). We now discuss observations that establish not only that SS433 has radio jets, but that they are linearly polarized with changes in strength, position, polarization, and polarization angle on time scales of 1–2 weeks.

Table 1 SS433 total flux densities at different frequencies and spectral indices

Date	JD–2444000	1,465 MHz S_ν (Jy)	2,695 MHz S_ν^* (Jy)	4,885 MHz S_ν (Jy)	8,085 MHz S_ν^* (Jy)	15,035 MHz S_ν (Jy)	α
8 September 1979	125	0.86 ± 0.1		0.41 ± 0.05		0.23 ± 0.05	–0.61
9 September 1979	126			0.41 ± 0.05			
16 September 1979	133			0.45 ± 0.05			
21 September 1979	138	1.0 ± 0.1	0.67 ± 0.06	0.45 ± 0.05	0.33 ± 0.05	0.22 ± 0.05	–0.64
7 December 1979	215	1.75 ± 0.1	1.37 ± 0.13	0.88 ± 0.05	0.65 ± 0.06	0.40 ± 0.05	–0.76
7 March 1980	306	0.62 ± 0.1	0.59 ± 0.06	0.41 ± 0.05	0.31 ± 0.05		–0.57
5 April 1980	335	0.78 ± 0.05	0.55 ± 0.05	0.43 ± 0.05	0.38 ± 0.05	0.20 ± 0.05	–0.54
20 June 1980	411	1.00 ± 0.05	0.70 ± 0.07	0.50 ± 0.05	0.28 ± 0.05	0.20 ± 0.05	–1.0

* Green Bank interferometer data (from ref. 12).

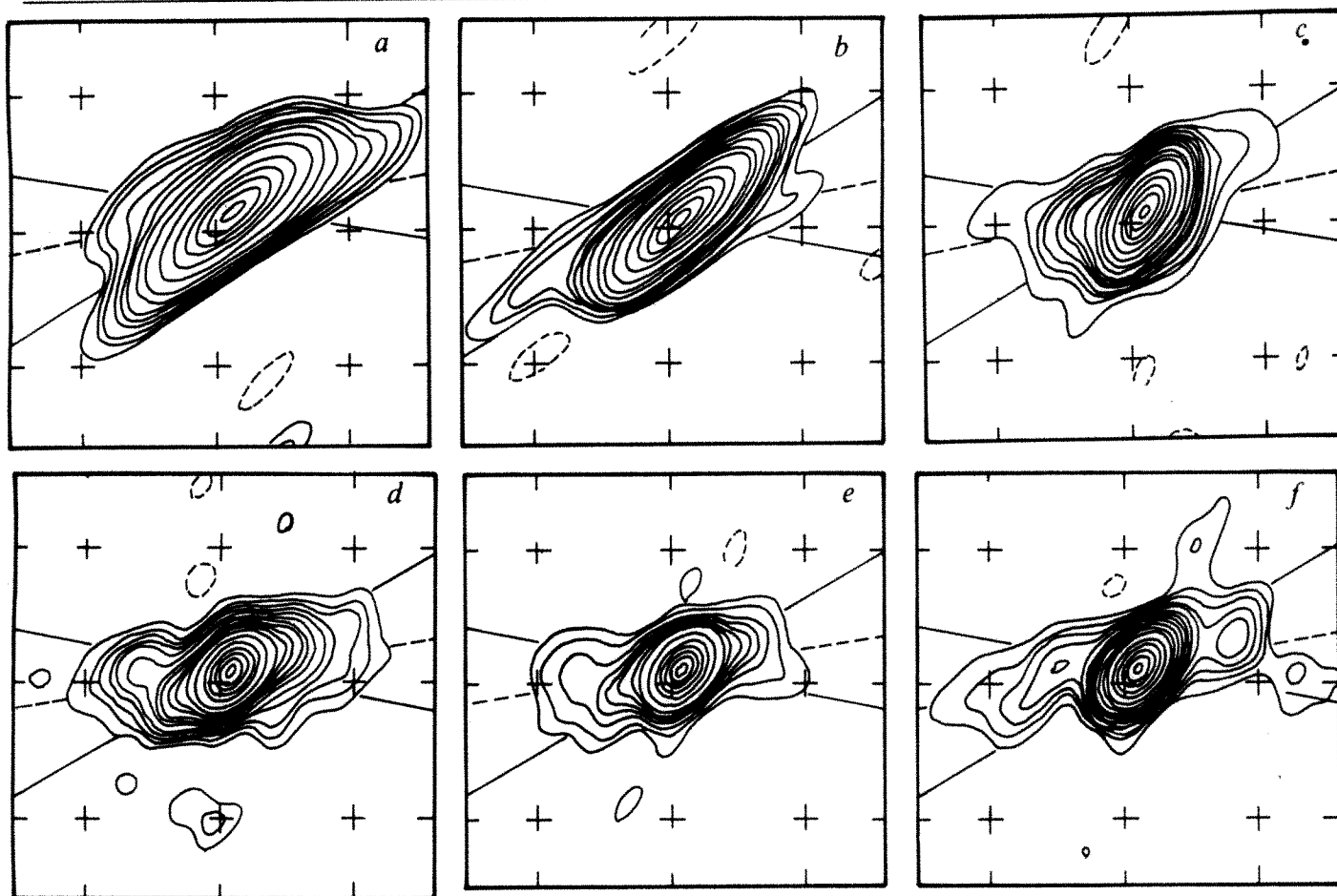


Fig. 1 VLA maps at 4,885 MHz. The grid of crosses corresponds to 2 arc s intervals. Position angles corresponding to 80°, 100°, and 120° are marked with solid, dashed, and solid straight lines, respectively. The beam shape corresponds to the 50% contour level for each map: *a*, 8–9 September 1979 with peak of 0.32 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 12, 8, 5, 4, 3, 2, 1.5, 1, and –1% level; *b*, 16 September 1979 with peak of 0.36 Jy per beam and contour levels as for the 8–9 September map; *c*, 7 December 1979 with peak of 0.71 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 15, 10, 7, 5, 4, 3, 2, 1.5, 1, 0.5, and –0.5% level; *d*, 7 March 1980 with peak of 0.29 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 12, 9, 7, 5, 4, 3, 2, 1, 0.7, 0.4, and –0.4% level; *e*, 5 April 1980 with peak of 0.34 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 10, 7, 5, 4, 3, 2, 1, 0.5, and –0.5% level; *f*, 20 June 1980 with peak of 0.39 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 15, 10, 7, 5, 4, 3, 2, 1, 0.7, 0.4, and –0.4% level. The restoring beam is denoted by the cross-hatched ellipse and is equivalent to the 50% contour of the map.

The observations

The dates of observation, frequencies, and total flux densities (S_ν) obtained for SS433 with the partially completed VLA from 8 September 1979 to 20 June 1980 are listed in Table 1, together with the 2,695 and 8,085-MHz flux densities for SS433 measured at the same times¹² with the Green Bank interferometer. The VLA observations were taken with different numbers of antennas, ranging from 17 in September 1979 to 22 in June 1980. Bandwidths of 50 MHz were used and system temperatures were roughly 45, 45, and 250 K at 1,465, 4,885, and 15,035 MHz, respectively. Calibration was based on interleaved observations of the point source 1947+079, assuming a position for this source of $\alpha(1950) = 19^h 47^m 40.15^s \pm 0.013^s$ and $\delta(1950) = +07^\circ 59' 35.8'' \pm 0.2''$ and flux densities of 1.15, 1.04, and 0.45 Jy at 1,465, 4,885, and 15,035 MHz, respectively. All the VLA observations listed in Table 1 are consistent with the following position for SS433: $\alpha(1950) = 19^h 09^m 21.274^s \pm 0.013^s$, $\delta(1950) = +04^\circ 53' 54.3'' \pm 0.2''$. This is consistent with the radio and optical position of Kaplan *et al.*¹³ measured near epoch 1979.6. Thus from July 1979 (epoch 1979.6) through June 1980 (epoch 1980.5), the radio position of SS433 has remained centred on the position of the optical image measured at epoch 1979.5.

The radio spectra for SS433 at six different epochs as established by the VLA and Green Bank interferometer data (Table 1) show non-thermal spectral indices of $\alpha =$

$-0.6 \pm 0.1 (S_\nu = \nu^\alpha)$ except for times shortly after flaring, that is, 7 December 1979 and 20 June 1980. Shortly after strong flaring the spectra steepen due to newly accelerated relativistic electrons. On 7 December, 7 March and 20 June there are signs of a turnover at lower frequencies.

The data obtained on 8 and 9 September 1979, when two observing programmes coincidentally included SS433 for different reasons, showed no signs of change. These data were combined in aperture synthesis maps to give radio maps for SS433 on 8–9 September 1979, 16 September 1979, 7 December 1979, 7 March 1980, 5 April 1980, and 20 June 1980. The data obtained on 21 September 1979 were at only 1 h angle and did not yield a useful map. On the last four epochs maps were made at 1,465, 4,885, and 15,035 MHz. Figures 1–3 show the 4,885-MHz aperture synthesis maps for six epochs, in which the

Table 2 Angular distance (ϕ) of identifiable polarized components from SS433

Date	JD–2444000	ϕ				
		A	B	B'	C	C'
7 December 1979	215	0.5"				
7 March 1980	306	1.4"	0.7"	0.6"	0.1"	
5 April 1980	335	1.6"	1.0"	0.8"	0.3"	
20 June 1980	411	2.2"	1.7"	1.5"	0.9"	0.8"

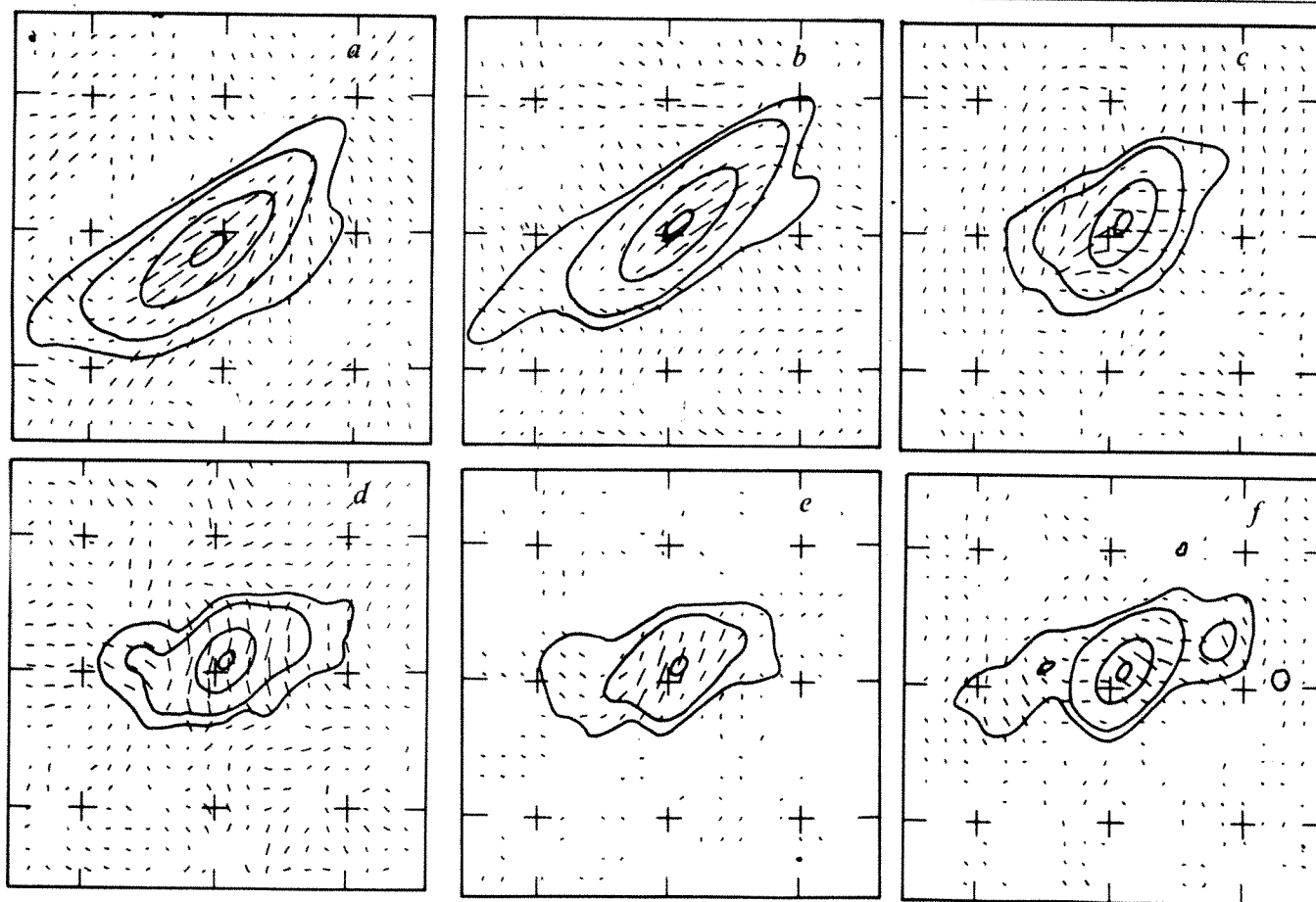


Fig. 2 VLA 4,885-MHz linear polarization vector maps superimposed on total intensity contours selected from Fig. 3. Polarization maxima correspond to the longest vectors with values of: a, 4.6 mJy for 8–9 September 1979; b, 7.6 mJy for 16 September 1979; c, 4.3 mJy for 7 December 1979; d, 2.6 mJy for 7 March 1980; e, 4.3 mJy for 5 April 1980; and f, 3.2 mJy for 20 June 1980.

grid of crosses corresponds to 2-arc s intervals. Figure 1 shows the total intensity maps (Stokes parameter I). Figure 2 displays the linear polarization vectors superimposed on a few total intensity map contours. Figure 3 shows maps of linear polarized intensity. In Figs 1 and 3 the lowest contour levels correspond to the peak 'noise' features in the maps. The noise level in linear polarization is indicated in Fig. 2 by the random vectors outside the vicinity of the SS433 radio contours.

The 4,885-MHz maps in Fig. 1 were processed with the standard mapping and sidelobe correction (CLEAN¹⁴) programmes used at the VLA site. In addition, a self-calibration technique, based on solving for antenna-based phase differences and amplitude ratios between the visibility data and a model derived from the Högbom¹⁴ CLEAN algorithm, was used to improve the dynamic range of the maps. This technique was used for all of the maps presented here. The 4,885-MHz maps in Fig. 1 show that the total intensity structure is predominantly contained within the arc set by the position angles 80°–120°, with a mean value of 100°. The VLA maps produced by Stocke *et al.* (in refs 15, 16) are qualitatively similar. These maps are dominated by a strong central component which usually contains ~75% of the total flux. Noting that the beam shape is identical to the 50% contour levels, there is a substantial

improvement in the maps as the VLA evolved from 14 to 22 usable antennas. The extended structure in these maps is below the 5% contour level. The improved synthesized beam and sensitivity of the VLA made the December 1979 maps greatly superior to the September 1979 maps, and the three maps made in 1980 had even better quality. The September 1979 maps showed signs of extended structure. The 7 December 1979 map was made during a short lull in an unusually strong flaring period that occurred from 7 August to 29 December 1979 (ref. 12). On 7 December 1979 there was twice the total radio flux of the other maps in Fig. 1, and the 'extra' flux appeared as a combination of more flux in the central source and a strong bulge of emission to the east which contains twice the flux of other extended features for the maps at other epochs. This bulge is probably due to plasma ejected during this intense period of flaring as it takes ~80 days for plasma at a speed of 0.26c to travel 1 arc s at a distance of 3.5 kpc. As the maps in Fig. 1 show drastic changes in structure, the time scale for perceptible changes in extended structure is almost certainly 1–2 weeks.

Radio jets on both the eastern and western sides of the central radio source are most easily seen in Fig. 1d–f. On 7 March most of the extended emission to the east shows a variation of position angle with distance from the central source. The same pattern continues through the April and June 1980 maps, with distinct changes in the structure from one map to the next. However, outward motion and the consequent determination of proper motions are best seen in the linear polarization maps of the SS433 radio jets. In the 7 March and 5 April maps in Fig. 1 the extended structure can be described in terms of three types of features: (1) extended emission located on both sides of the central source with roughly 3 arc s separation and a position

Table 3 Proper motion of SS433 polarized components

A	(V/d) sin θ (arc s per day)		C
	B	B'	
0.0081	0.0094	0.0089	0.0078

Table 4 Observed flux density and polarization parameters of SS433

JD - 2444000	Frequency (MHz)	Core flux density (Jy)	Extended flux density (Jy)	Peak polarized flux density (Jy)	Total polarized flux density (Jy)	Fractional polarization (%)
125-126	4,885	0.25	0.16	0.0046	0.014 ± 0.004	9
133	4,885	0.34	0.11	0.0076	0.023 ± 0.004	21
215	4,885	0.64	0.24	0.0043	0.019 ± 0.003	8
306	4,885	0.27	0.14	0.0026	0.023 ± 0.002	16
335	4,885	0.32	0.11	0.0043	0.017 ± 0.002	16
411	4,885	0.39	0.11	0.0032	0.017 ± 0.002	16
335	1,465	0.53	0.25	0.0025	0.038 ± 0.003	15
411	1,465	0.83	0.20	0.0025	0.032 ± 0.003	16

angle of 88° , (2) low level emission with size scales up to 6 arc s, with a full range of position angles between 80° and 120° , and (3) a clearly identifiable outer extrusion on the east side with position angle near 80° . In the 20 June map there is only the central source and extended structure evolved from expansion of the structure seen in March and April.

In Fig. 2 the 4,885-MHz polarization vectors are superimposed on a few total intensity contours. In all cases the polarization of the total flux is of the order of 1% or less, but much higher percentages (8–20%) of polarization are seen in the extended features of the jets. This is most clearly seen in the later four maps in Fig. 3, where the linearly polarized intensity maps are shown. In Fig. 3 X shows the location of the central radio source coincident with the optical image of SS433. The very one-sided linear polarization map for 7 December clearly indicates a result supported by succeeding maps—newly produced structures of relativistic plasma are highly linearly polarized once they are more than a few light days from SS433.

Because of this, the linear polarized intensity maps are both indicators of ordered magnetic field regions, and good indicators of the variable 'jet' structure of SS433.

The changes in polarization angles are very striking in Fig. 2. Figure 3 shows that in the September and December maps the polarization is dominantly one sided. This was during a period of heavy flaring activity¹². Between 7 March and 12 June 1980, no strong flaring occurred¹² but there was an onset of flaring on 13 June 1980. During this period the polarized emission appears decidedly two-sided with respect to the star. On 20 June 1980 the double structure has become quite extended, but a new component has appeared to the west of SS433. The changes in polarization can be understood if it is realized that the polarization vectors in Fig. 2 reflect the extended structure rather than the central radio source, in which the polarization averages out to very low levels. In the 7 December 1979 map the polarization angle rotates by nearly 90° from the inner side of the ejected material to the outer side indicating that the magnetic

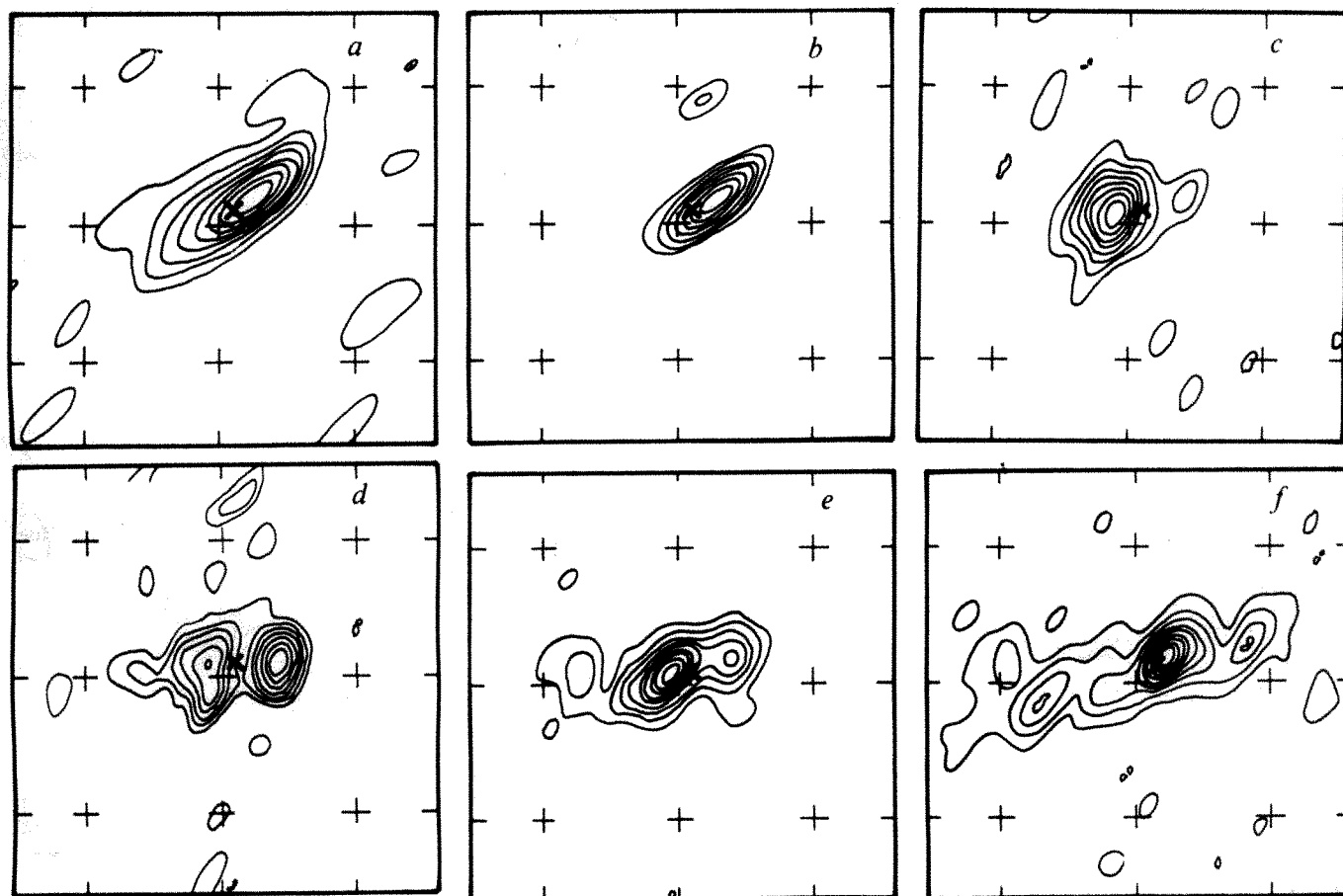


Fig. 3 VLA 4,885-MHz linearly polarized intensity maps with contours at 10% levels starting with 90% and peak values in Jy per beam and lowest contour levels of: a, 4.6 mJy + 20% 8–9 September 1979; b, 7.6 mJy + 30% 16 September 1979; c, 4.3 mJy + 20% 7 December 1979; d, 2.6 mJy + 30% 7 March 1980; e, 4.3 mJy + 10% 5 April 1980; f, 3.2 mJy + 10% 20 June 1980. The grid of crosses corresponds to 2 arc s intervals.

fields are being rotated parallel to the flow of ejected material. This follows from the fact¹⁷ that in the transparent part of the spectrum the polarization of a synchrotron radiation source is perpendicular to the magnetic field, and for a uniform source with parallel field lines the fractional polarization is $(3\gamma+3)/(3\gamma+7)$, where the electron energy distribution is $KE^{-\gamma}$. As the typical spectral index of -0.6 corresponds to an electron energy distribution with $\gamma = 2.2$, fractional polarization of uniform field regions would be expected to be 0.71. Considering that the observed fractional polarization of the eastern features in the 7 December, 7 March, 5 April and 20 June maps is 8, 16, 16, and 16% respectively, there are clearly field inhomogeneities which would be expected from the complex geometries involved. In the 7 March map, where eastern and western jets have roughly the same flux levels, the inner polarization angles of the polarized double are roughly 0° (to the north) and are consistent with magnetic fields parallel to the inner flow regions. The strength of the outer feature to the east in Fig. 3d is too weak to consider the apparent rotation of the polarization angle to be real. However, the excellent signal-to-noise ratio of the 5 April polarization map shows that the polarized double source has both expanded and rotated its position angle to $\sim -25^\circ$. The signal-to-noise level of the eastern polarization feature suggests that the average magnetic field direction is perpendicular to the flow of radiating relativistic particles. In the polarized source map of 20 June the outermost features are the evolved versions of the double of the 7 March and 5 April maps with polarization angles rotated to -150° . The feature associated with newly appearing polarized flux in the 20 June map shows an ejected

flow at a position angle near 120° and a polarization angle of $\sim -60^\circ$, so again the magnetic field lines are nearly parallel to the flow.

The previous discussion unfortunately assumes negligible Faraday rotation between SS433 and the Earth. Before any definitive information can be derived from the position angle information, multi-frequency observations of SS433 and background radio sources must determine the line of sight contribution to Faraday rotation.

Figure 4 shows the total intensity maps, the polarization vectors superimposed on a few total intensity contours, and the polarized intensity maps for 5 April and 20 June 1980. The dominant 1,465-MHz emission is more extended and may have a steeper spectrum than the features dominating the 4,885-MHz maps. However, the poor resolution of the 20-cm maps precludes useful determination of the spectral indices for anything except the total source. The polarized intensity maps clearly show two strong peaks separated by 2.3 and 3.0 arc s on 5 April and 20 June respectively. Because of poorer resolution, all of the polarized emission of the western and eastern 'jets' is coalesced into a simple polarized double source expanding with an apparent separation speed of $0.2c$. Noting the location of the components of the separating double in Fig. 3, most of this motion is clearly due to the western component. The linear polarization vectors of the western and eastern components of the double are nearly perpendicular to each other. As the polarization angle for the dominant eastern and western components in the 4,885-MHz maps is essentially the same, the apparent rotation between western and eastern features in the

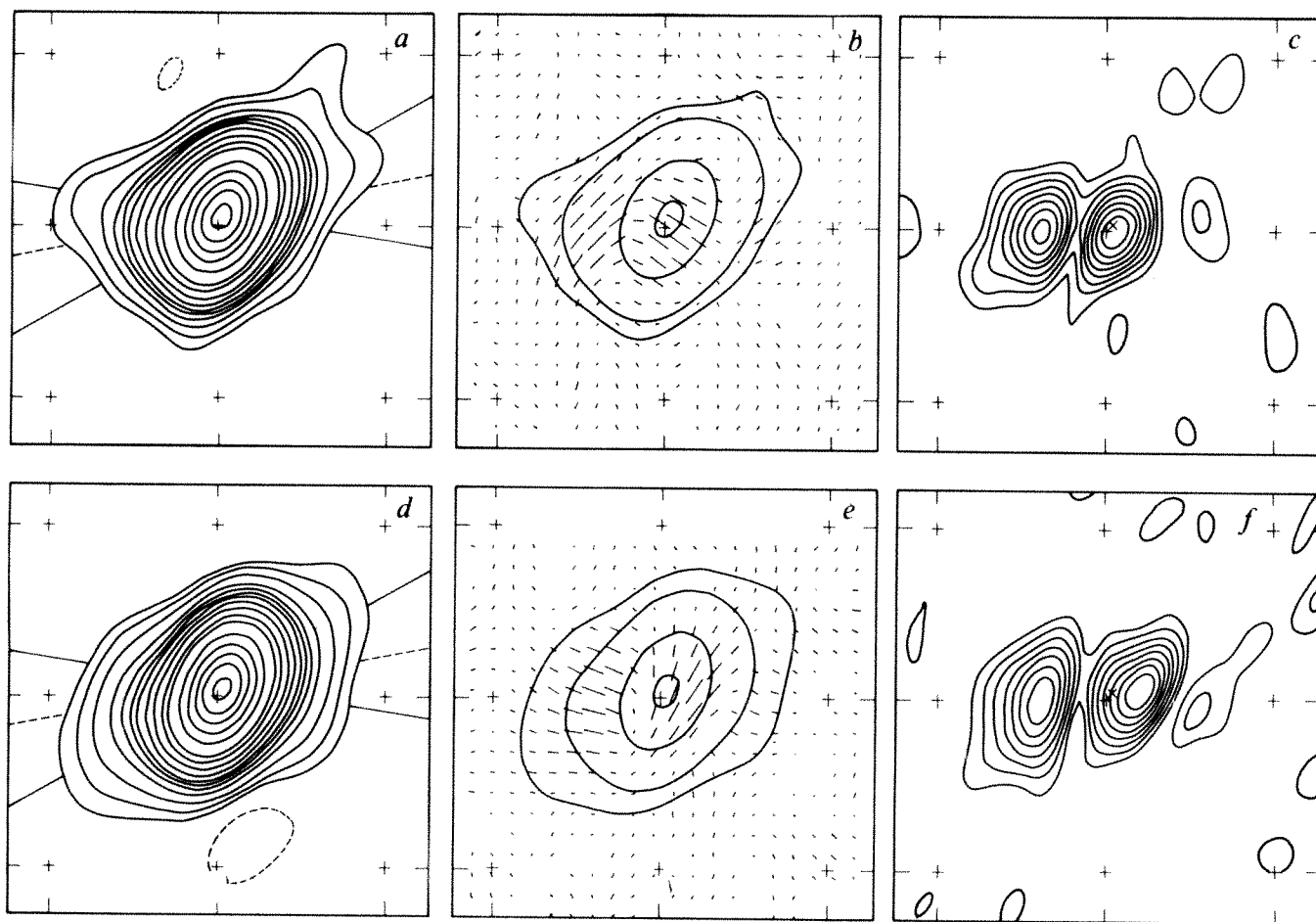


Fig. 4 VLA 1,465-MHz maps in total intensity, linear polarization vectors superimposed on selected total intensity contours, and linearly polarized intensity for 5 April 1980 in *a*, *b*, and *c*, respectively, and for 20 June 1980 in *d*, *e*, and *f*, respectively. The grid of crosses corresponds to 5 arc s intervals. Solid, dashed, and solid lines in *a* and *d* indicate position angles of 83° , 100° , and 117° respectively. In *a* the peak is at 0.63 Jy per beam while in *d* the peak is at 0.9 Jy per beam. In *a* and *d* the contours are at the 95, 80, 65, 50, 35, 20, 15, 10, 7, 5, 4, 3, 2, 1, 0.5, and -0.5% level, with *d* having lower contours at the $\pm 0.25\%$ level. In *b*, *c*, *e*, and *f* the peak polarization is 2.5 mJy per beam. In *c* and *f* the linearly polarized intensity contours are at the 90, 80, 70, 60, 50, 40, 30, 20, and -20% level. X, the location of the central radio source and optical image.

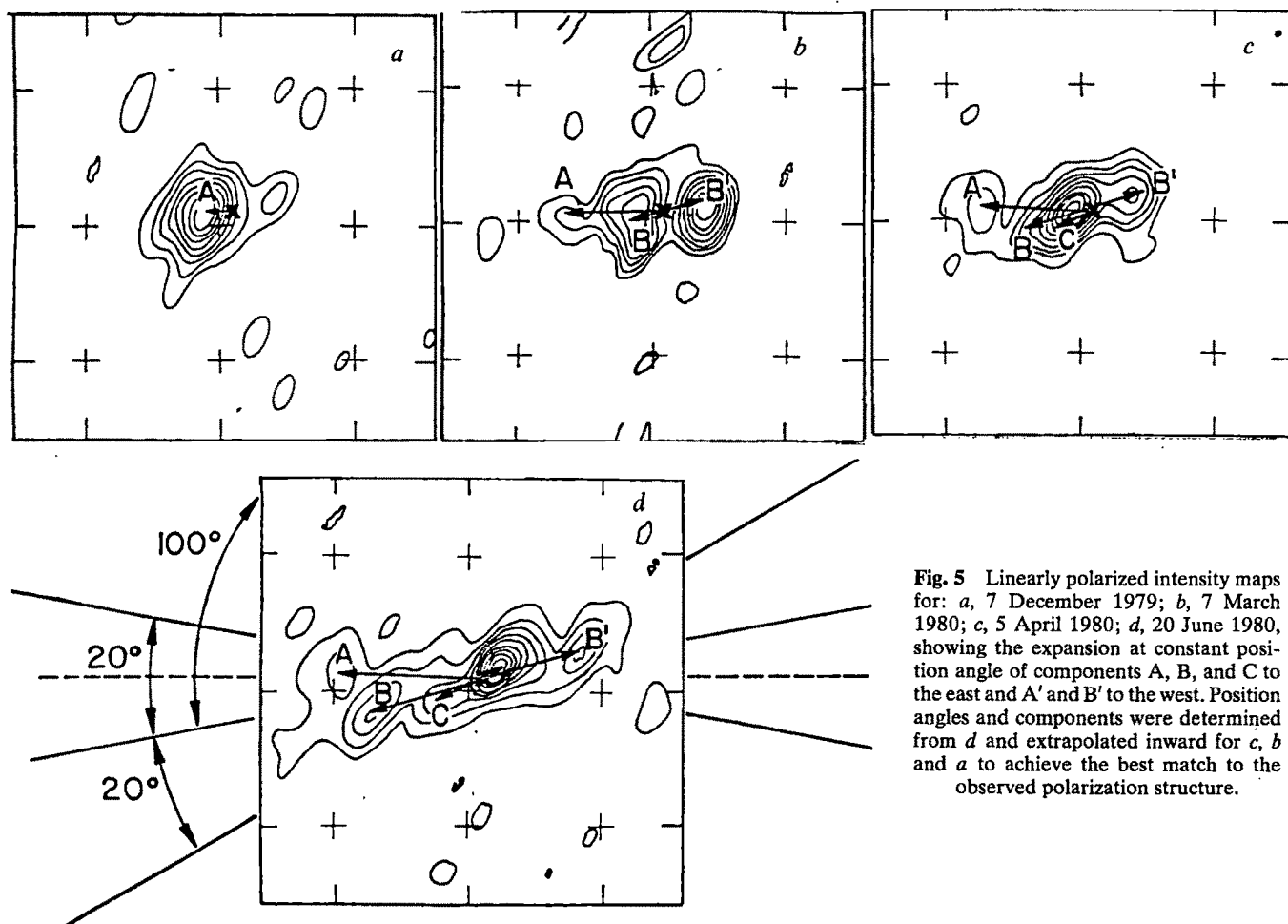


Fig. 5 Linearly polarized intensity maps for: *a*, 7 December 1979; *b*, 7 March 1980; *c*, 5 April 1980; *d*, 20 June 1980, showing the expansion at constant position angle of components A, B, and C to the east and A' and B' to the west. Position angles and components were determined from *d* and extrapolated inward for *c*, *b* and *a* to achieve the best match to the observed polarization structure.

1,465-MHz maps may be due to Faraday rotation produced by a plasma region around SS433.

The 15,035-MHz maps corresponding to the entries in Table 1 showed a dominant point source at the position of SS433, with only marginal signs of extended structure that were too close to the noise level to be considered significant.

Transverse motions of SS433 jet features

The most obvious indication of transverse motions in the extended structure of SS433 is the changing separation of the double source in the 1,465-MHz polarized maps. Portions of the S-shaped structures can be seen that would be expected from ejection of material at different time with different ejected position angles. This is most easily traceable in the later four maps in Figs 1 and 3. However, it is difficult to identify the locations of moving features. Fortunately, the low degree of polarization of the dominant central radio source and the relatively high degree of polarization of the extended structure means that the maps of linearly polarized intensity are a good means of evaluating the changes of extended structure with time. The maps made on 7 December, 7 March, 5 April and 20 June, 1980 are nearly ideal for tracing the evolution of identifiable features.

Figure 5 shows the linearly polarized emission for 7 December, 7 March, 5 April and 20 June 1980, with possible components identified as A, B, and C on the east side and A' and B' on the west side. The identification of features and the determination of position angles were made solely from the 20 June 1980 polarized map. The position angles for the A, BB', and C features are $88^\circ \pm 5^\circ$, $108^\circ \pm 5^\circ$, and $115^\circ \pm 10^\circ$ respectively. The vectors for the corresponding features in the 7 December, 5 April and 7 March maps were drawn preserving the 20 June position angles, but shortening the vectors to best achieve a match with the earlier polarized intensity maps. Table 2 shows the resulting angular distances (ϕ) for each

feature in each map. The uncertainties for the angular displacements in Table 2 are at least 0.1 arcs, and can be much higher when merged or nearly merged features are hard to locate. These angular displacements as a function of time are linear and their slopes—the proper motion of the features—are listed in Table 3. The average proper motion of the features is 0.0088 arc s per day.

As the features labelled A, B, and C have the most clearly identifiable motion, they can be used to identify when features moving at a proper motion of 0.0088 arc s per day would coincide with SS433. The results are JD - 2444000 = 139, 217, and 297, for A, BB', and C, respectively. From coincidence of probable origin time and position angle, the very weak features A and BB' in the 1980 linear polarization maps are almost certainly the remnant of the strong flaring that occurred between 7 August and 29 December 1979. Weakening of these features with time is not surprising as one of the main physical processes affecting the relativistic plasma in the SS433 jets is adiabatic expansion and energy loss.

The proper motion of the features is $(v/d) \sin \theta$, where v is the velocity of the feature, d is the source distance, and θ is the angle the feature is moving with respect to the observer. Using the ephemeris of Margon *et al.*¹⁵

$$\cos \theta = \pm[0.34 \cos 2\pi(\psi + \alpha) + 0.167] \quad (1)$$

where $\psi = (\text{JD} - 2433555.6 \text{ days})/164.0 \text{ days}$, $\alpha = 0.330$, the inclination angle of the system with respect to the sky is 80° and the precessing beam angle is 20° . Obtaining θ from the times of ejection gives for A, BB', and C 114° , 90° , and 109° . These angles yield a mean proper motion v/d of 0.0088 arc s per day. Because little or no slowing down of motion of the features is observed, their velocity is probably $0.26c$. This results in a distance to SS433 of 5.1 kpc.

It is difficult to explain the individual structure of the polarized emission of the ejected features. The A feature is decidedly one sided. The time of ejection of this feature corresponds to

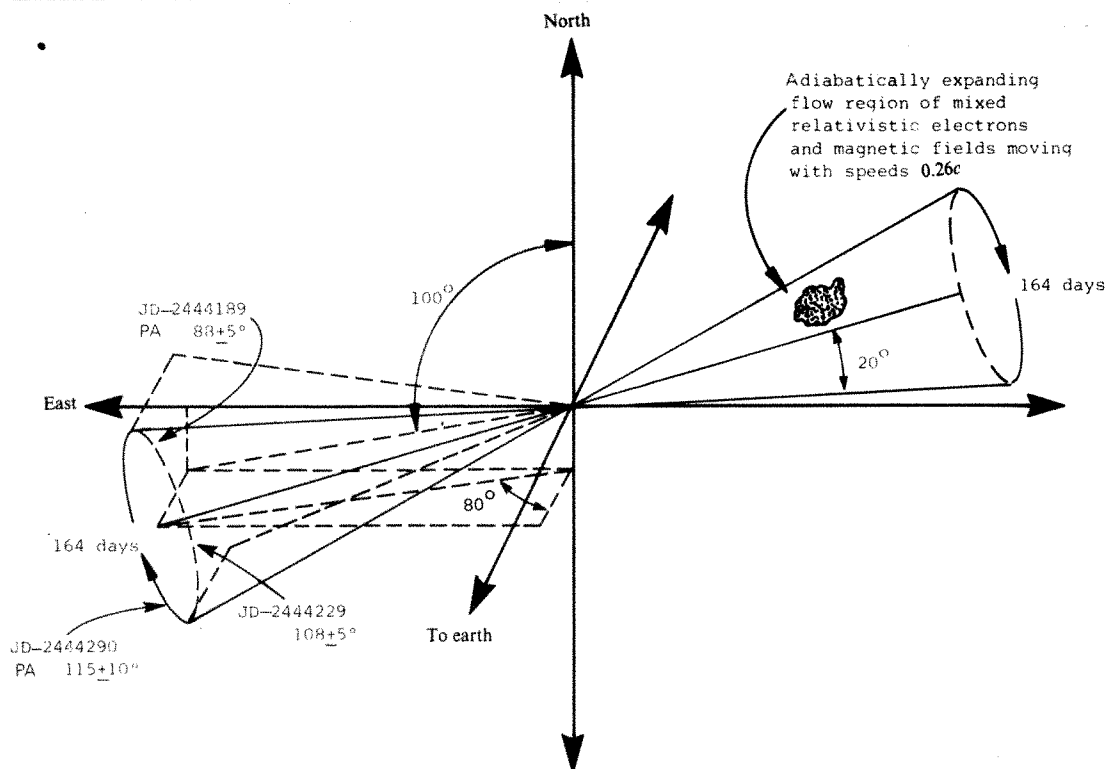


Fig. 6 A schematic diagram of the kinematics of the SS433 system with some parameters as determined by Margon¹⁵, but with the 100° apparent mean position angle as determined from the radio data and identification of foreground/background location on the basis of the strength bias for eastern jets and the tendency for western polarized features and western jets to appear suddenly at distances of a few tenths of an arc s from the compact radio source. The location of the identifiable features A, B, and C are indicated along with the values of position angle and JD-2444000 at the time the observed transverse velocities would make them coincident with SS433.

individual flaring events in the Green Bank data¹². The time of the ejection of the BB' feature corresponds to a period of intense complex flaring such that multiple flares were ejected, perhaps along the two jets. The C feature is very weak and was ejected during a quiescent period.

The apparent motion of the BB' features indicates the same ejection angles on both sides of SS433, and thus ejection at the same epoch is likely. For features ejected simultaneously along the two beams, the proper motions of the two features must be corrected for time travel effects of special relativity. This effect allows the velocity of the jets to be distinguished unambiguously; the value of v/d will then give the distance to SS433. Unfortunately, the BB' ejection occurred at an angle θ close to 90°, thus negating any time travel effects between the BB' jets.

Optical jets and changing radio structure

The most important fact relating the optical jets of SS433 and the changing radio structures with angular size scales of 0.1–5 arc s is the coincidence of the $\pm 20^\circ$ cone angle for the apparent precession of the optical jets¹⁵ and the $100^\circ \pm 20^\circ$ region in which features are ejected with different position angles at different times. The other important data determined by the VLA maps, that should be related to the optical jets, are the fairly well-defined position angles of $88^\circ \pm 5^\circ$, $108^\circ \pm 5^\circ$, and $115^\circ \pm 15^\circ$ for the A, BB', and C features presumably ejected from the central regions of SS433 at JD-2444000 = 139, 217, and 297, respectively. Figure 6 indicates the location of these position angles on one of the possible geometries for the SS433 optical and radio jets. The association of the direction of rotation with the eastern jet is arbitrary. The western jet is put in the background because its features seem to appear suddenly at a distance of 2–3 light days from SS433. There is a strong bias for western jets to be weaker, and the observed Faraday rotation may indicate the presence of an absorbing region near the centre that affects the inner western jets more than the inner eastern jets. These choices, however, do not affect the important points concerning the correlation of radio and optical jets. Assuming the beam to precess in 164 days, we note from Fig. 6 that there is reasonable consistency between the initial position angles for the A, BB', and C features, their date of ejection, and their relation to a 164-day period. Based on this geometry, the radio data, and a 164-day period, the maximum red/blue shift point would be predicted to occur on JD2444170 and the maximum blue/red

shift point would be predicted to occur on JD2444250, with uncertainties of the order of 10–20 days. According to Margon¹⁵ the maximum red/blue shift and blue/red shift points were reached on JD2444157 and JD2444239. Considering the uncertainties in position angles and average transverse speeds, we conclude that the behaviour of the moving features in the VLA maps is consistent with an association of optical jets and radio jets in the inner regions where the flows producing the radio emitting particles originate. Without more detailed tracking of identifiable features, probably with maps made at weekly intervals, a better correlation of extended radio jets and optical jets cannot be established from VLA data alone.

Parameters of the SS433 radio source

When synchrotron radio sources have known distances, sizes, flux densities, and observable Faraday rotation, several parameters of the radio source can be estimated, albeit in a model-dependent manner. We now consider typical values of the core radio source and typical values of one of the extended jet structures.

Using a combination of surface integrals over appropriate portions of the total intensity and linearly polarized intensity maps, the flux densities and polarization characteristics as listed in Table 4 can be determined. In Table 4 the core flux density was determined from the sum of the clean components found before any extended components were subtracted. Thus the core sources in Table 4 are located on top of the plateau level of the extended structure. The extended flux density was determined by the difference between the total integrated flux and the core flux density. The fractional polarization was determined by the ratio of polarized flux density to extended flux density.

The best maps of total intensity and polarized intensity are for JD-2444000 = 306, 335, and 411. The fractional polarization for the extended structure for these maps is surprisingly close to a single value, 16%, for both 1,465 and 4,885 MHz. However, as this was a period with relatively little flaring, it is hard to assess the significance of this polarization compared with the more variable values made with earlier, poorer VLA maps, though the 8% polarization determined for the 7 December 1979 extended structure should be of comparable quality to later results.

Based on Table 4 we adopt the following values as typical for the SS433 central source and an extended lobe defined as one half of the extended structure: core source of 0.3 Jy at

4,885 MHz, with a size of 0.1 arc s (7.5×10^{15} cm assuming a distance of 5.1 kpc); and extended lobes of 0.05 Jy at 4,885 MHz, with a size scale of 1 arc s (3×10^{16} cm). One may calculate the magnetic field (B) and the relativistic electron energy (W_e) assuming equipartition of energy and magnetic fields^{18,19} which are $W_e = 1.5 \times 10^{42}$ erg and 3.1×10^{42} erg, $B = 0.0077$ and 0.0014 G, respectively, for the core source and extended lobe. These magnetic fields indicate synchrotron loss lifetimes of 10^9 – 10^{11} s. Therefore synchrotron losses cannot have a significant role in the evolution of the observed central and extended radio emission.

If we assume that the 90° polarization angle rotation between the two lobes of the 20 cm extended emission is due to Faraday rotation, we estimate a minimum value of electron concentration of 3 – 4 electrons cm^{-3} in regions with the product size and magnetic field as derived above.

The above approximations will have less effect on the results than the real changes of parameters due to known changes in the radio source, with the possible exception of an erroneous assumption that energy in relativistic protons can be neglected in the equipartition of energy argument used to determine the magnetic field.

Conclusions

In addition to highly variable radio emission coincident in position with the optical image of SS433, there are highly polarized radio jets with size scales 0.1 – 5 arc s. The extended structure that we describe as jets for convenience is confined to position angles of $100^\circ \pm 20^\circ$. Features occurring in this structure, particularly in maps of linearly polarized flux, can be traced over time periods of 200 days with a proper motion of

0.0088 arc s per day with constant position angle. Extrapolating back to a hypothetical time of ejection based on the observed proper motion, the position angles of the radio jet features are consistent with those expected from the optical jet locations at that time. If the velocity of the polarized features is the same as that of the optical emission lines, the source distance is 5.1 kpc.

The 100° position angle is very close to the position angle determined by the long baseline observations of Spencer²⁰ and the VLBI observations of Walker *et al.*²¹. It is also very close to the position angle for the bulges of the large scale W50 radio source surrounding²² SS433. The $100^\circ \pm 20^\circ$ cone of relativistic plasma ejection defined by the radio maps corresponds reasonably well to the position angles of the large scale X-ray emission seen by Seward *et al.*²³.

The estimated magnetic fields are 0.001 – 0.01 G, so synchrotron loss lifetimes are $>10^9$ s; therefore synchrotron losses are negligible compared with the adiabatic losses associated with the expansion of newly created relativistic plasma components.

The evolution of the radio structure of SS433 indicates that the basic model must be that of adiabatically expanding flows which are accelerating relativistic particles and generating magnetic fields while moving into a turbulent medium made up from previous ejections. Fitting a more complicated model to the variable structure of this source is left to a future analysis of these and more extensive VLA observations. The capability to map evolving structures that change with time scales of the order of a week is one of the unique advantages of studying SS433 with the VLA.

The VLA and the National Radio Astronomy Observatory are operated by Associated Universities, Incorporated, under contract with the NSF.

Received 14 November 1980; accepted 13 January 1981.

1. Abell, G. P. & Margon, B. *Nature* **279**, 701–703 (1979).
2. Liebert, J., Angel, J. R. P., Hege, E. K., Martin, P. G. & Blair, W. P. *Nature* **279**, 384–387 (1979).
3. Margon, B. *et al. Astrophys. J. Lett.* **230**, L41–L45 (1979).
4. Margon, B., Ford, H. C., Grandi, S. A. & Stone, R. P. S. *Astrophys. J. Lett.* **233**, L63–L68 (1979).
5. Fabian, A. & Rees, M. J. *Mon. Not. R. astr. Soc.* **187**, 13P–16P (1979).
6. Milgrom, M. *Astr. Astrophys.* **78**, L9–L12 (1979).
7. Begelman, M. C., Sarazin, C. L., Hatchett, S. P., McKee, C. F. & Arons, J. *Astrophys. J.* (in the press).
8. Marshall, F. E., Swank, J. H., Boldt, E. A., Holt, S. & Serlemitsos, P. J. *Astrophys. J. Lett.* **230**, L145–L148 (1979).
9. Ryle, M., Caswell, J. L., Hine, G. & Shakeshaft, J. *Nature* **276**, 571–575 (1979).
10. Seaquist, E. R., Garrison, R. E., Gregory, P. C., Taylor, A. R. & Crane, P. C. *Astr. J.* **84**, 1037–1041 (1979).

11. Wade, C. M. & Hjellming, R. M. *Astrophys. J.* **170**, 523–528 (1971).
12. Johnston, K. J. *et al. Astr. J.* (submitted).
13. Kaplan, G. H., Kallarkal, V. V., Harrington, R. S., Johnston, K. J. & Spencer, J. H. *Astr. J.* **85**, 64–65 (1980).
14. Högbom, J. A. *Astr. Astrophys. Suppl.* **15**, 417–422 (1974).
15. Margon, B. *Scient. Am.* **243**, 54–65 (1980).
16. Gilmore, W. & Seaquist, E. R. *Astr. J.* **85**, 1486–1495 (1980).
17. Kellermann, K. I. *Galactic and Extra-Galactic Radio Astronomy* Ch. 12 (eds Verschuur, G. L. & Kellermann, K. I.) 320–352 (Springer, New York, 1974).
18. Ginzburg, V. L. & Syrovatskii, S. I. *Ann. Rev. Astr. Astrophys.* **3**, 297–350 (1965).
19. Hjellming, R. M. *Physics and Astrophysics of Neutron Stars and Black Holes* (eds Giacconi, R. & Ruffini, R.) 185–188 (North-Holland, 1978).
20. Spencer, R. E. *Nature* **282**, 483–484 (1979).
21. Walker, R. C. *et al. Astrophys. J. Lett.* (in the press).
22. Geldzahler, B., Pauls, T. & Salter, C. *Astr. Astrophys.* **84**, 237–244 (1980).
23. Seward, F., Grindlay, J., Seaquist, E. & Gilmore, W. *Nature* **287**, 806 (1980).

Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur

Wayne A. Hendrickson

Laboratory for the Structure of Matter, Naval Research Laboratory 6030, Washington DC 20375, USA

Martha M. Teeter

Department of Chemistry, Boston University, 685 Commonwealth Avenue, Boston, Massachusetts 02215, USA

The highly ordered crystal structure of crambin has been solved at 1.5 Å resolution directly from the diffraction data of a native crystal at a wavelength remote from the sulphur absorption edge. The molecule has three disulphide bridges among its 46 amino acid residues, of which 46% are in helices and 17% are in a β -sheet. Crambin is shown to be an amphipathic protein, inasmuch as its six charged groups are segregated from hydrophobic surface elements. Phasing methods used here will also apply elsewhere.

BIJVOET'S observation¹ that anomalous scattering might aid in solving the phase problem has had many implications², including suggestions that departures from Friedel's law of

diffraction symmetry might suffice to determine directly the atomic structures of crystals containing heavy atoms^{3–5}. Anomalous-scattering methods have also been widely used

Table 1 Expected diffraction ratios for potential applications of resolved anomalous phasing

Molecule	N_P	N_A	$\langle \Delta F \rangle / \langle F_P \rangle$	$\langle F_A \rangle / \langle F_P \rangle$
Sulphur-rich proteins:				
Crambin	400	6S	1.4%	29%
Snake neurotoxin ³⁸	500	8S	1.5%	30%
Metalloproteins:				
Hi PIP ³⁹	700	4Fe/8S	5.3%	38%
Haemerythrin ¹⁹	1,000n	2nFe	3.0%*	17%*
Oligonucleotides:				
d(CpGpCpGpCpG) ⁴⁰	300	10P	1.6%	39%
Oligopeptides:				
Antamanide ⁴¹	100	10O	0.5%†	37%
Heavy-atom complexes:				
Myoglobin ⁴²	1,300	1Hg	4.5%	31%
Lysozyme ⁴³	1,100	1U	8.5%	39%

N_P is the approximate number of non-hydrogen atoms in the total molecule and N_A is the number of anomalous scatterers. (Omission of hydrogen atoms has a negligible effect on these calculations: $\langle F_P \rangle$ with hydrogens is $\sim 1.01\langle F_P \rangle$ without hydrogens.) The diffraction ratios are estimates for zero scattering angle. In the case of only one kind of anomalous scatterer these ratios are $\langle \Delta F \rangle / \langle F_P \rangle \approx 2^{1/2} (N_A^{1/2} \Delta f_A') / (N_P^{1/2} Z_{\text{eff}})$ and $\langle F_A \rangle / \langle F_P \rangle \approx (N_A^{1/2} Z_A') / (N_P^{1/2} Z_{\text{eff}})$ where Z_{eff} is the effective atomic number (~ 6.7 for non-hydrogen protein atoms) and Z_A' is $Z_A + \Delta f_A'$. F_P and F_A are structure-factor magnitudes of real contributions from the complete molecule and from anomalous scatterers only, respectively. Of course, the diffraction ratios will vary with scattering angle in the same way as the ratio of thermal and scattering factors for the anomalous scatterers to those of the other atoms. Except where noted all estimates are for CuK α radiation.

* At low resolution (for example, 5.5 Å) these values are enhanced by a factor of $2^{1/2}$ because the dimeric iron centre is then unresolved.

† Oxygen anomalous scattering is computed for CrK α radiation.

in protein crystallography⁶⁻¹¹, particularly as an adjunct to isomorphous-replacement phasing. We have now combined elements from these two traditions to solve the structure of crambin by exploiting the anomalous scattering of sulphur atoms at a single wavelength (1.54 Å of CuK α) far removed from the absorption edge of sulphur (5.02 Å).

Crambin is a small protein found in the embryonic tissue (cotyledons and hypocotyledons) of seeds from *Crambe abyssinica*, a relative of mustard and rape commonly known as Abyssinian cabbage. It is hydrophobic in that organic solvents are required to solubilize it. Van Etten *et al.*¹² characterized crambin after noticing that crystals formed during evaporation of an aqueous acetone extract of defatted seed meal. The remarkable crystalline order observed by Teeter for this protein (strong diffraction from spacings < 0.88 Å)¹³ correlates well with the unprecedented structural stability seen in solution by Llinás *et al.*¹⁴ using NMR spectroscopy. The function of crambin is not known. However, the recently completed chemical sequence¹⁵ reveals an unmistakable homology with the plant toxins purothionin¹⁶ and viscotoxin¹⁷. Crambin itself is not toxic when fed to rats¹⁸.

We set out to solve the crystal structure of crambin in view of its exceptional potential for providing detailed structural information and in the hope of shedding light on the function of this hydrophobic protein. When our efforts to prepare heavy-atom derivatives failed, we explored alternatives to isomorphous replacement for phase determination. Our experience in using anomalous scattering to locate the iron atoms in myohaemerythrin¹⁹ led us to contemplate an attempt based on the six sulphur atoms in crambin. An estimate of the magnitude of Bijvoet differences to be expected from crambin gave a value about half that found with myohaemerythrin and showed that sulphur scattering would account for 98% of the expected anomalous signal. The expected contribution of the sulphur partial structure to the protein diffraction pattern proved to be 29% ($\langle |F_S| \rangle / \langle |F_P| \rangle$), or 9% based on $|F|^2$ ratios. This

seemed to provide a plausible, albeit marginal, basis for the crystal structure analysis.

The amino acid sequence determination, begun to aid the crystallographic analysis, was about 65% complete when the first electron-density maps were obtained. Completion of the chemical sequencing was aided by, proceeded alongside and helped to verify the structure solution. The sequence, as reported elsewhere¹⁵, is shown in Fig. 1.

Diffraction measurements

The diffraction data were measured from a single crambin crystal, grown and mounted as described before¹³, in equilibrium with its mother liquor of 60% aqueous ethanol (v/v). This crystal had approximate dimensions of $0.2 \times 0.5 \times 0.4$ mm and gave unit cell parameters of $a = 40.96$, $b = 18.65$, $c = 22.52$ Å and $\beta = 90.77^\circ$. The space group is P2₁. X-ray reflections from Ni-filtered CuK α radiation were recorded at room temperature (22–24 °C) as peak-top ω step-scans on a Picker FACS-1 diffractometer that is controlled by the Vanderbilt program system²⁰. For spacings greater than 1.5 Å, both I_h and I_{-h} were measured. Care was taken to reduce systematic errors in the Bijvoet differences by successively collecting blocks of 25 reflections at $(\phi, \chi, 2\theta)$ and then their Friedel counterparts at $(\phi, \chi, -2\theta)$. After completing the 1.5 Å data set, $+2\theta$ measurements were continued in shells out to 0.95 Å spacings. At this point the crystal slipped and further measurements were abandoned.

The scans for the 1.5 Å data comprised 4-s counts at each of seven steps in 0.03° intervals in ω and at a single background point. Counts were corrected for possible coincidence losses. These data were then reduced to integrated intensities using the procedure of Hanson *et al.*²¹, which fits gaussian profiles over smoothly varying backgrounds. Standard corrections were made for Lorentz and polarization factors, absorption²² and scale²³. No correction was made for radiation damage. (A monitor reflection showed only 3% decay in I after 150 h exposure.) Residual systematic errors in the Bijvoet differences

$$\Delta F \equiv |F_h| - |F_{-h}| \quad (1)$$

were minimized by applying local scale factors²⁴ parameterized to be anisotropic. These scale factors, applied to $|F_{-h}|$, ranged from 0.97 to 1.01 and eliminated errors that swamp the signal. After scaling, r.m.s. (ΔF) is 2.3e for acentric data as compared with 1.4e for the centric ($h0l$) data where there is no signal. This symmetry estimate or error agrees well with that obtained from counting statistics and instrumental instability, r.m.s. ($\sigma_{\Delta F}$) = 1.6e. Thus, on average the anomalous-scattering signal is about 1.7e, $(\langle (\Delta F)^2 \rangle - \langle \sigma_{\Delta F}^2 \rangle)^{1/2}$, which is 2.1% of the r.m.s. (F).

This work concerns only the 1.5 Å data. The data from beyond 1.5 Å have not been processed.

Sulphur structure

The first task of the analysis, to locate the sulphur atoms, was accomplished by interpreting Patterson maps computed with coefficients of $(\Delta F)^2$ as suggested by Rossmann⁸. Peaks in such maps correspond to inter-atomic vectors between anomalous scatterers. However, because the coefficients involve small differences, the maps are very sensitive to erroneous measurements. Hence, differences with $|\Delta F| > 5$ r.m.s. (ΔF) were rejected as distribution outliers (four reflections) and the more

Thr - Thr - Cys - Cys - Pro₅ - Ser - Ile - Val - Ala - Arg₂₀
 - Ser - Asn - Phe - Asn - Val₁₅ - Cys - Arg - Leu - Pro - Gly₂₀
 - Thr_{Pro} - Glu - Ala - Ile - Cys - Ala - Thr - Tyr - Thr_{Ser}
 - Gly - Cys - Ile - Ile - Ile₃₅ - Pro - Gly - Ala - Thr - Cys₂₀
 - Pro - Gly - Asp - Tyr - Ala₂₀ - Asn

Fig. 1 The amino acid sequence of crambin. Details of the sequence analysis are given in ref. 15.

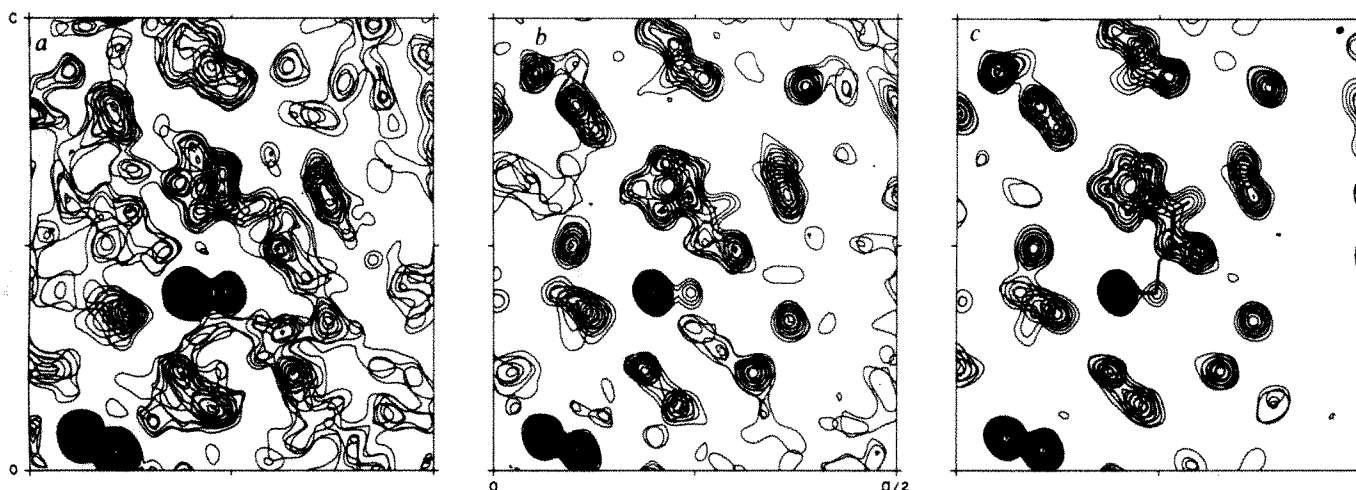


Fig. 2 A portion of the electron-density distribution for crambin at three stages in the analysis. Each frame encloses an area bounded by $0 < x < a/2$ and $0 < z < c$ and is a composite of the sections at $y = (19-21)b/40$. The adjacent, dense spherical features at the lower left are the sulphur atoms of Cys 16–Cys 26. The dense spheroid near the centre of each frame is S^* of Cys 32. S^* of Cys 4 lies just to the right of this and is centred above these sections. The ring of Pro 5 is seen in the plane of these sections. All the maps are at 1.5 Å resolution. Frame *a* is part of the initial Fourier synthesis from which the structure was interpreted. Coefficients of $m(\frac{1}{2}(|F_h| + |F_{-h}|) \exp(i\phi))$ were used with figures of merit and phases determined as follows. Bimodal anomalous distributions for which the partial-structure choice-probability²⁹ exceeded 0.7 were given the phase of the chosen mode and weighted by the intrinsic figure of merit²⁹ (30% of the 5,660 reflections). Phasing for relatively sharp unimodal distributions, $m_{\text{ano}} > 0.4$ and $m_{\text{ano}} > m_{\text{combine}}$, was kept at ϕ_{ano} and m_{ano} (8%). For the remaining cases, the centroid phases and weights, m_{combine} , were determined from multiplicatively combined distributions³⁰. These included the centric zonals (11%) and error rejects (2%), for which there was no anomalous phasing, as well as the other general reflections (48%). Partial structure probabilities were computed with $Q = 2$. Overall, $\bar{m} = 0.55$. Frame *b* is from the first $(2F_{\text{obs}} - F_{\text{calc}})$, ϕ_{calc} map after direct refinement of the initial model to an agreement factor of $R = 0.322$. Frame *c* is the latest $(2F_{\text{obs}} - F_{\text{calc}})$, ϕ_{calc} map from the model at $R = 0.104$.

error-prone weak data with $|F_h|$ or $|F_{-h}| < 5\sigma_F$ were also excluded (127 reflections). Small Bijvoet differences, $|\Delta F| < 1.2\sigma_{\Delta F}$, were also left out (3,038 reflections) but this had little impact on the maps.

These anomalous-difference Patterson maps were remarkably clean. The first map was computed at 3 Å resolution and could readily be interpreted as arising from three sites that correspond to unresolved disulphide units. When the data to 1.5 Å spacings were also included, the disulphide Patterson peaks became resolved into clusters of peaks from which the six-atom sulphur structure was deduced.

It is easy to show⁹⁻¹¹ that for a structure containing relatively weak anomalous scatterers all of one kind, for example, sulphur in crambin,

$$\Delta F \approx -2\delta \sin(\psi - \phi) \quad (2)$$

Here, $F \exp(i\phi)$, where $F \approx \frac{1}{2}(|F_h| + |F_{-h}|)$, is a structure factor from the real parts of the scattering and $\delta \exp(i\psi)$ is calculated from the anomalous scatterers when given the imaginary components, $\Delta f''$, for scattering factors. This relationship, which also explains the effectiveness of the $(\Delta F)^2$ Patterson synthesis, provides a basis for refining the structure of anomalous scatterers and ultimately for phasing the whole structure. As $|\Delta F|$ becomes large the sine factor must approach ± 1 so that $|\Delta F| \approx 2\delta$. Thus, atomic parameters of anomalous scatterers can be refined against $|\Delta F|$ values if only the largest differences are included. This was first done for myohaemerythrin to give a quite precise Fe–Fe distance¹⁹.

Refinement of the crambin sulphur structure was against a data set restricted, by rejection criteria described above, to the 1,910 strongest from among a total of 5,017 Bijvoet differences. The value of $\Delta f''$ for sulphur was taken to be 0.557 (ref. 25). The refinement reduced R to 0.33. Disulphide distances, unrestrained, in the resulting structure were in good agreement with expectation: $d = 2.00, 2.02$ and 2.04 Å with $\sigma_d = 0.06$ Å.

Phase determination from partial-structure resolved anomalous scattering

Phase information from anomalous scattering at a single wavelength, like that from a single isomorphous replacement

experiment, is intrinsically ambiguous. Given an observed ΔF and the values of δ and ψ calculated from refined atomic parameters, equation (2) can be solved for the desired phase, ϕ , for the reflection. However, the possible solutions at $\phi = \psi + \pi/2 \pm \theta$, $\theta = \cos^{-1}(\Delta F/2\delta)$, are equally likely. Moreover, due to the substantial errors in ΔF , other phases may also be plausible or, quite possibly, a formal solution will be precluded. This ambiguity and imprecision in phases can be taken into account by using a probability treatment of the kind introduced by Blow and Crick²⁶. We have used an error model¹¹ whereby the probability distribution, $P_{\text{ano}}(\phi)$, for phase information from anomalous scattering is

$$P_{\text{ano}}(\phi) = N \exp \{ -[\Delta F + 2\delta \sin(\psi - \phi)]^2 / 2E^2 \} \quad (3)$$

The standard error, $E = (\sigma_{\Delta F}^2 + E_0^2)^{1/2}$, is composed of the error

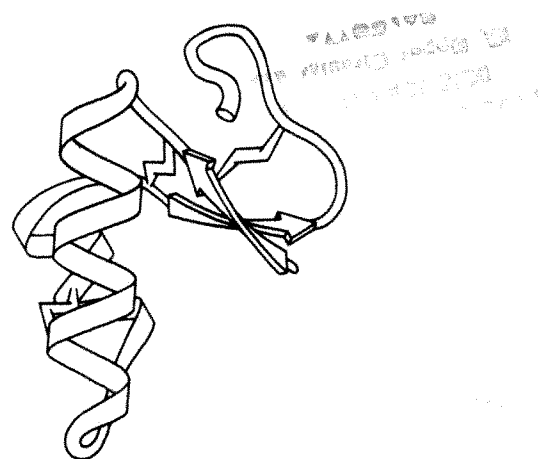


Fig. 3 A schematic drawing of the backbone of crambin. This representation of crambin was drawn by Jane Richardson to be faithful to computer drawings of the skeleton. Arrows depict β -strands. The disulphides are drawn as 'lightning flashes'. The reader looks down along *b*; *a** runs from left to right and *c* runs from page bottom to page top.

in ΔF and a residual lack-of-closure error, E_0 (ref. 27). N is a normalization factor.

The phase ambiguities, $\pm\theta$, must be resolved for the anomalous scattering to be of use. If the anomalous scatterers are sufficiently heavy, then phases for the total structure, ϕ , will tend to be close to those for the heavy-atom structure, ψ . One can then simply take the anomalous alternative that is nearer to the heavy-atom phase³⁻⁵. However, this process is often not valid if, as in the case of crambin, the partial structure of anomalous scatterers is not a dominating influence. Thus, we have used a probability treatment where the distribution, $P_{\text{par}}(\phi)$, for phase information for the sulphur partial structure is based on that given by Sim²⁸

$$P_{\text{par}}(\phi) = N' \exp \{2Q|F_p||F_s| \cos(\psi - \phi) / \langle F_U^2 \rangle\} \quad (4)$$

Here, $\langle F_U^2 \rangle$ is the expected value of the scattering contribution from the unknown part of the structure; $|F_p|$ and $|F_s|$ are the structure factor moduli observed for the protein and calculated from the sulphur structure, respectively; Q is an arbitrary sharpening factor.

These probability representations were used to combine the phase information from anomalous scattering with that from the

sulphur partial structure and also to compute figures-of-merit, m , for weighting the Fourier synthesis. After some experimentation we settled on a scheme for probabilistic choices²⁹. A choice was made if the partial-structure probabilities (equation (4)) discriminated well between alternative maxima in the anomalous scattering distribution (equation (2)). Sharp unimodal distributions were used directly. Otherwise, $P_{\text{ano}}(\phi)$ and $P_{\text{par}}(\phi)$ were multiplicatively combined³⁰.

There remains one further ambiguity. The sulphur structure deduced from Patterson maps could equally well be of either hand. Nothing in the anomalous-scattering and partial-structure information decides directly between enantiomers, but resulting phases will only be sensible when based on the correct choice. Independent information is required. Chemical reasonableness in a Fourier synthesis seemed to favour one sulphur enantiomer. This choice was corroborated by Sayre phase refinements³¹ on both hands. (This 'refinement' reduced the residuals to $R = 0.23$, but it actually degraded the map.)

Interpretation of the electron-density map

A part of the Fourier synthesis in the proper hand is shown in Fig. 2a. The most striking density in this map comes from the

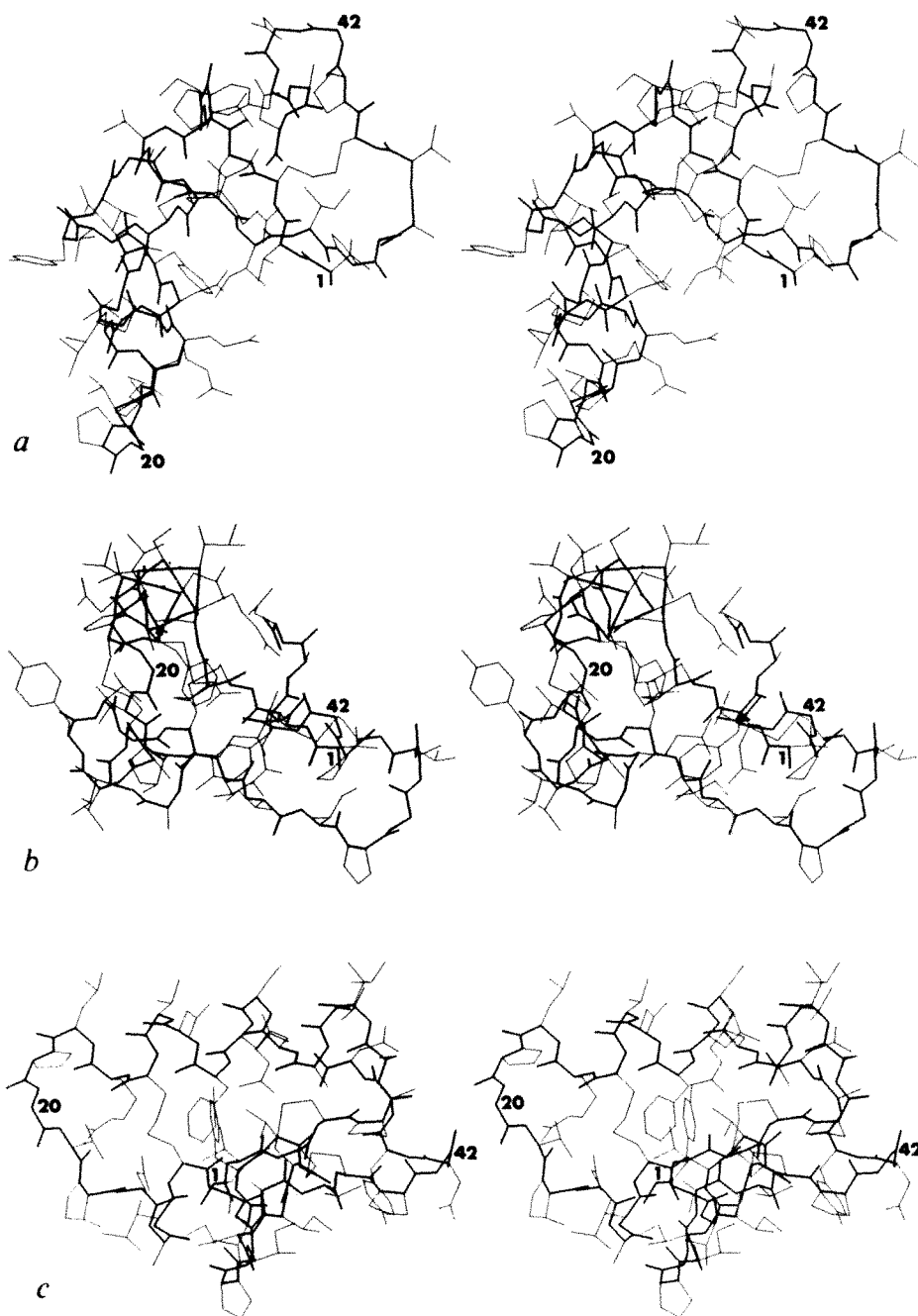


Fig. 4 Stereo drawings of the crambin atomic structures. Bonds along the main-chain backbone are enhanced. The conformational and compositional heterogeneity at positions 7, 22 and 25 is included. Residue numbers are given near the C^α positions of Thr 1, Gly 20 and Gly 42 in each of the views. *a*, View, as in Fig. 3, through the thinnest dimension. *b*, View along *c* with *a** running from left to right. This is approximately along the helical stem and perpendicular to the β -sheet (towards the back at the right). *c*, View into *a** with *c* running from left to right. This is approximately perpendicular to the helix axes as seen from the β -arm side. These stick figures and the surface diagrams (Figs 5 and 6) were all drawn using the program PLT1 written by G. J. Quigley.

known sulphur atoms, but other rather atomistic features are also evident. Yet, the map has quite an uneven quality. In an initial survey we correctly identified four side groups (later to be Pro 5, Tyr 44, Phe 13 and Asn 14), but peptide linkages to adjacent residues were not readily apparent even with knowledge of a tentative sequence for the first 33 residues. However, it did prove possible to fit the prominent proline ring (Fig. 2) and two sulphur atoms with a model of the unique -Cys 3-Cys 4-Pro 5-peptide, albeit through some weak density.

The seeming reasonableness of these parts of the map encouraged us to attempt a complete interpretation. The fitting was made without recourse to graphics sophistication. Rather, we simply followed density features contoured on paper ($5 \text{ mm } \text{\AA}^{-1}$), proceeding outward from the disulphide bridges while bearing in mind standard stereochemical rules. Putative atomic positions were marked on the sheets. Several polypeptide segments evolved and eventually joined up. The last connections, residues 18-22 and 37-41, were rather tenuous but they did complete the chain. It was a heartening verification of the model when helices were finally recognized. Atomic coordinates were read off the map with the aid of a gridwork overlay.

Model refinement and revision

The initial model included 338 non-hydrogen atoms from 45 amino acid residues and 18 water molecules. This model fitted the diffraction pattern with a reliability index $R = \Sigma |F_{\text{obs}} - F_{\text{calc}}| / \Sigma F_{\text{obs}}$ of 0.43 for the data from spacings between 5 and $1.5 \text{ \AA}$. Twelve cycles of stereochemically restrained refinement³² reduced R to 0.32 while imposing moderately good geometry as typified by the r.m.s. deviation from bond ideality of $0.04 \text{ \AA}$. Individual isotropic thermal parameters were varied but with tight restraints³³.

Several iterations of map interpretation, model revision and continued refinement followed. Inspection of the first ($2F_o - F_c$) synthesis (Fig. 2b) and the corresponding ($F - F_c$) difference map revealed a number of misplaced atoms, particularly in segments 18-22 and 37-41. A model with the polypeptide backbone revised in these segments then refined to an R of 0.27. Repair of side chains (40% were wrongly identified initially in the unsequenced C-terminal portion), addition of water mole-

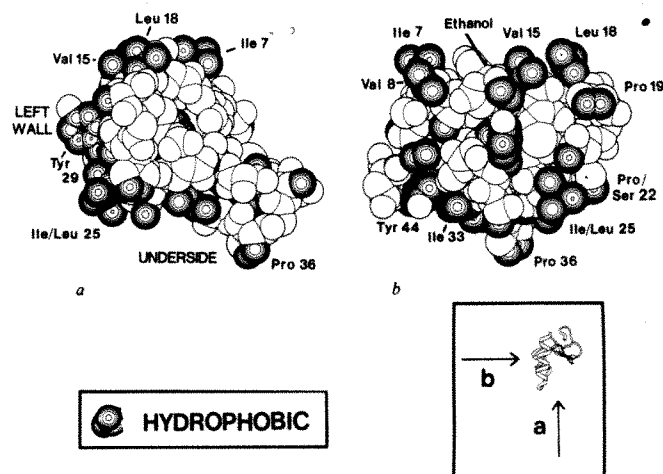


Fig. 6 Hydrophobic surface of crambin. The radii used here are the same as those in Fig. 5. Atoms designated as hydrophobic here only include the side-chain carbon atoms of Ala, Val, Leu, Ile, Phe and Tyr residues, C^β and C^γ of Pro and C^γ of Thr. *a*, View looking into the hydrophilic inner bend, as in Figs 4b and 5b, showing hydrophobic residues that line the left-wall and underside surfaces. *b*, View looking into the left-wall surface along a^* with c running from right to left. Note that the polar side chain of Asn 12 is partially covered by the hydrophobic cloak of an ethanol molecule.

cules to a total of 39, and further refinement brought R down to 0.17. It then became apparent that some of the supposed water sites actually constituted a 46th residue.

Finally, special attention was paid to solvent regions and to side-group assignments. Another ~40 solvent molecules were located on inclusion of the $5\text{--}10 \text{ \AA}$ shell of data and, after release of thermal restraints, residue identifications were completed down to the distinctions between N and O needed to assign Asp 43 and Asn 46. This took R to 0.12. Side-group assignments were confirmed at this point by the completed chemical sequence analysis. A model was then devised to account for the conformational and compositional heterogeneity at three residue positions. Isotropic refinement of this model (414 atomic sites including 72 water and 4 ethanol molecules) against all 5,638 data in the $10\text{--}1.5 \text{ \AA}$ shell gave $R = 0.114$ with bond ideality of $0.018 \text{ \AA}$. Anisotropic refinement with restraints³³ such that the r.m.s. bond-distance fluctuation was kept to $0.05 \text{ \AA}$ gave $R = 0.104$ (see Fig. 2c). The resulting coordinates will be deposited in the Protein Data Bank.

Molecular conformation

Crambin has the shape of the Greek capital letter gamma (Γ) when viewed as in Fig. 3. (We describe crambin in terms of a model so oriented lying on a table before the reader.) The stem of the Γ is an antiparallel pair of helices and the cross-arm consists of two antiparallel β -strands, an irregular strand and a classic β -turn. The backbone skeleton of crambin (Fig. 3) gives an impression that the relatively thin β -arm is only tenuously connected to the helical stem. This belies the exceptional degree of rigidity in crambin¹⁴. In fact, side chains fill the axillary juncture, or inner bend, between the stem and the arm with numerous stabilizing salt bridges, hydrogen bonds and other contacts (Figs 4, 5). The disulphide bridges further fix the internal structure of the sub-domains.

The local conformational parameters of crambin fit remarkably well with accepted principles of secondary structure³⁴. Backbone conformation angles are all within allowed regions of the (ϕ, ψ) plot, although the conformations at three glycine residues (20, 31 and 37) would be disallowed for any other residue. All the peptide bonds are *trans*. The present restrained model has an r.m.s. deviation of 3.8° from peptide planarity. Only the ω angles following residues 29-31 and 44-45 deviate

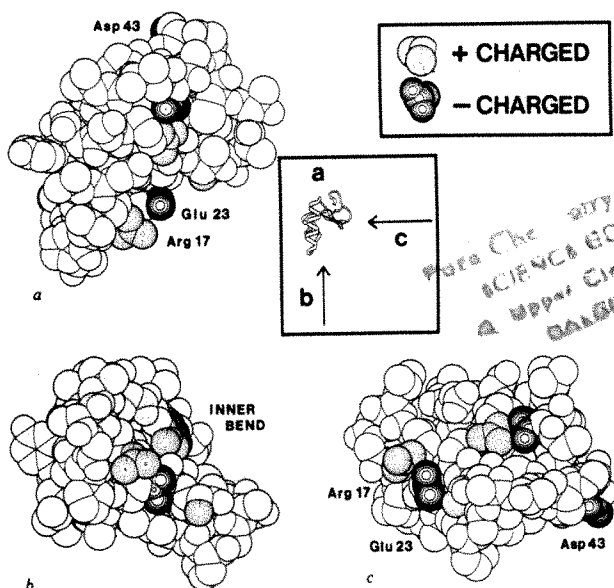


Fig. 5 Ionic surface of crambin. These surface representations of crambin have the following Van der Waals radii: C, $1.7 \text{ \AA}$; N, $1.4 \text{ \AA}$; O, $1.4 \text{ \AA}$ and S, $2.1 \text{ \AA}$. No hydrogen atoms are included. Positively charged groups of atoms are shaded with dotted circles and negatively charged groups are shaded with solid circles. *a*, View of the molecule as in Fig. 3 and in Fig. 4a. *b*, View looking into the inner bend of the Γ from along the helix axes as in Fig. 4b. *c*, View looking into the inner bend perpendicular to the helix axes as in Fig. 4c.

from planarity by more than 5°, the greatest deviation being 11°. Side-chain conformation angles also lie close to 'ideal' values. The 43 χ angles that are expected to be staggered have an r.m.s. deviation of only 9.5°. Two of the three aromatic residues are within 4° of the expected transverse conformation at χ_2 , but Tyr 29 has $\chi_2 = 55^\circ$. The disulphide torsion angles, which are also expected to favour $\pm 90^\circ$, are -79° (3-40), 106° (4-32) and -86° (16-26).

Residues 1-4 and 32-35 produce β -strands linked by four hydrogen bonds. The (ϕ, ψ) angles in these segments average $(-121^\circ, 148^\circ)$. This two-stranded β -sheet has the usual left-handed twist when viewed perpendicular to the strands. Residues in extended conformations at 39-41 form a third strand that is not in the β -sheet.

Residues 7-19 and 23-30 are in helices. The first 10 residues in the upper helix have (ϕ, ψ) angles that average $(-62^\circ, 42^\circ)$ and only vary by $\sigma_\phi = \sigma_\psi = 4^\circ$. A disruptive proline at the C-terminus of this helix is accommodated by 3_{10} -helix conformations at positions 17 and 19. The last residue in the lower helix, Thr 30, has a distorted 3_{10} -like conformation, but otherwise this helix is also α -helical. Its (ϕ, ψ) angles average $(-65^\circ, 37^\circ)$ but it is somewhat less regular than the upper α -helix, $\sigma_\phi = 6^\circ$ and $\sigma_\psi = 8^\circ$. The angle between helix axes is 139° .

Residues 5-6, 20-22, 31, 36-38 and 42-43 produce the five turns in the crambin tertiary fold. Four of these segments contain or are immediately preceded by proline residues and four include glycines. The first four turns are relatively gradual and each is negotiated in a characteristic way. However, the last turn is a sharp reversal, with a carbonyl-41 to amide-44 hydrogen bond. It is a type I β -turn, the most common kind, and has (ϕ, ψ) angles of $(-62^\circ, -23^\circ)_{42} \rightarrow (-90^\circ, -2^\circ)_{43}$.

Conformational heterogeneity exists at Ile 7 and Ile 25 where C δ^1 atoms take either of two staggered possibilities. Compositional heterogeneity also exists at positions 22 and 25. Refinement parameters suggest that residue 22 is ~60:40 Pro/Ser and that residue 25 is ~60:40 Ile/Leu. The heterogeneity at residue 22 apparently causes a disorder in Tyr 29; the refined position of its O ^{n} makes an impossibly short contact of 2.6 Å with C δ of Pro 22 on a screw-related molecule.

Amphipathic protein surface

The surface of crambin has an amphipathic character. The six charged groups (1 NH $^+$, 10 Arg $^+$, 17 Arg $^+$, 23 Glu $^-$, 43 Asp $^-$ and 46 COO $^-$) as well as several other hydrophilic side chains are clearly segregated from what otherwise is largely a hydrophobic molecular surface. The surface of crambin can be understood as comprising three components. The face of the inner bend between the helical stem and the β -arm is a sloping surface that is almost entirely hydrophilic (Fig. 5). The other two surface elements, namely the left wall of the helical stem and the underside of the molecule, are primarily hydrophobic (Fig. 6). The left-wall surface curves gently around the outer bend while the underside surface is relatively flat.

All but one of the charged groups are clustered on the face of the inner bend. Six of the 13 other polar side chains also line or rim this surface. These groups are highly interconnected. For example, the guanidinium group of Arg 10 makes a salt bridge to the terminal carboxyl oxygens of Asn 46, forms hydrogen bonds to the hydroxyl and carbonyl oxygens of Thr 2, and interacts with the amide oxygen of Asn 14 through a tightly bound water molecule. Bridging water molecules also link the carboxyl group of Glu 23 to the terminal amino group of Thr 1

and the guanidinium group of Arg 17. In the crystal this chain of alternating charges continues to the carboxyl group of Asp 43 in a molecule related by lattice translation.

Unlike the inner-bend surface, or the surfaces of water-soluble proteins in general, the remainder of crambin is interspersed with hydrophobic groups. The left-wall surface includes eight fully exposed leucine, isoleucine, valine, proline and tyrosine side groups and the underside surface contains seven such exposed groups. There are also exposed polar side groups on these surfaces, but several of these are hydrogen-bonded back to the main chain.

The most extensive packing contacts in the crystal involve Van der Waals interactions between juxtaposed left-wall surfaces. There are only four direct intermolecular hydrogen bonds in the asymmetric unit, all other lattice interactions being mediated through water molecules. Although water is excluded from the left-wall contacts, it surrounds the remaining available surface. Most of the water molecules are associated with the hydrophilic inner bend, but elsewhere a number are actually in contact with hydrophobic side chains. Surprisingly little of the ethanol is bound detectably. Only four ethanol sites with a total occupancy of 2.3 have been located, whereas the 72 identified water sites have occupancies totalling 52.7. If the solvent space (34% of the crystal) had the composition of the mother liquor (60% ethanol), we would expect 18 ethanol and 39 water molecules per asymmetric unit.

It is likely that the amphipathic character of the crambin surface has significance for the physiological activity of the molecule. Unfortunately, the biological role of crambin is unknown.

Prospective applications

It might seem that anomalous-scattering methods as used here would apply solely to special, highly ordered structures like crambin. However, estimates of phasing power suggest that these methods may have broad applicability. Expected diffraction ratios related to anomalous-scattering and partial-structure strengths have been calculated for various crystals of known structure (Table 1). Many of these problems give ratios that compare favourably with those for crambin and probably they too could have been solved from partial-structure resolved anomalous scattering. This is true both for structures with native anomalous scatterers and for heavy-atom complexes.

Although Bijvoet differences may be small, with appropriate experimental design most systematic errors can be eliminated. Moreover, as only a single crystalline species is used, derivatives need not be isomorphous. Also, phasing power carries on well to high resolution, although high resolution is not a requirement. We have recently solved the structure of a trimeric haemerythrin at 5.5 Å resolution using Fe-resolved anomalous phasing followed by molecular averaging for phase refinement (J. L. Smith and W. A. H., unpublished). The possibility of performing multi-wavelength experiments at synchrotrons³⁵⁻³⁷ enhances the promise for anomalous-scattering applications.

The crystal structure of crambin should be a good model for exploring protein conformations, fluctuations and solvent interactions in extraordinary detail. Further diffraction studies and toxicity analyses are in progress.

We thank Jane Richardson for her drawings, Gary Quigley for his plotting programs, and Janet Smith and Jerome Karle for discussions. This work was supported in part by NSF grant PCM80-03929 to M.M.T.

Received 14 October; accepted 23 December 1980.

1. Bijvoet, J. M. *Proc. Acad. Sci. Amst.* **B52**, 313-314 (1949).
2. Ramaseshan, S. & Abrahams, S. C. (eds) *Anomalous Scattering* (Munksgaard, Copenhagen, 1975).
3. Peerdeman, A. F. & Bijvoet, J. M. *Acta crystallogr.* **9**, 1012-1015 (1956).
4. Ramachandran, G. N. & Raman, S. *Curr. Sci.* **25**, 348-351 (1956).
5. Okaya, Y. & Pepinsky, R. *Phys. Rev.* **103**, 1645-1647 (1956).
6. Blundell, T. L. & Johnson, L. N. *Protein Crystallography* (Academic, London, 1976).
7. Blow, D. M. *Proc. R. Soc. A247*, 303-336 (1958).
8. Rossmann, M. G. *Acta crystallogr.* **14**, 383-388 (1961).
9. North, A. T. C. *Acta crystallogr.* **18**, 212-216 (1965).
10. Matthews, B. W. *Acta crystallogr.* **20**, 82-86 (1966).

11. Hendrickson, W. A. *Acta crystallogr.* **A35**, 245-247 (1979).
12. Van Etten, C. H., Nielsen, H. C. & Peters, J. E. *Phytochemistry* **4**, 467-473 (1965).
13. Teeter, M. M. & Hendrickson, W. A. *J. molec. Biol.* **127**, 219-223 (1979).
14. Llinás, M., De Marco, A. & Lecomte, J. T. J. *Biochemistry*, **19**, 1140-1145 (1980).
15. Teeter, M. M., Mazer, J. A. & L'Italien, J. J. *Biochemistry* (submitted).
16. Mak, A. S. & Jones, B. L. *Can. J. Biochem.* **22**, 835-842 (1976).
17. Samuelsson, G., Seger, L. & Olson, T. *Acta chem. Scand.* **22**, 2624-2642 (1968).
18. Van Etten, C. H. *et al. Cereal Chem.* **46**, 145-155 (1969).
19. Hendrickson, W. A., Klippenstein, G. L. & Ward, K. B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2160-2164 (1975).
20. Lenhart, P. G. *J. appl. Crystallogr.* **8**, 568-570 (1975).
21. Hanson, J. C., Watenpugh, K. D., Sieker, L. & Jensen, L. H. *Acta crystallogr.* **A35**, 616-621 (1979).

22. North, A. C. T., Phillips, D. C. & Mathews, F. S. *Acta crystallogr.* **A24**, 351–359 (1968).
23. Karle, J. & Hauptman, H. *Acta crystallogr.* **6**, 473–476 (1953).
24. Matthews, B. W. & Czerwinski, E. W. *Acta crystallogr.* **A31**, 480–487 (1975).
25. Cramer, D. T. in *International Tables of X-ray Crystallography* Vol. 4, 148–151 (Kynoch, Birmingham, 1974).
26. Blow, D. M. & Crick, F. H. C. *Acta crystallogr.* **12**, 794–802 (1959).
27. Ten Eyck, L. F. & Arnone, A. *J. molec. Biol.* **100**, 3–11 (1976).
28. Sim, G. A. *Acta crystallogr.* **12**, 813–815 (1959).
29. Hendrickson, W. A. *Acta crystallogr.* **B27**, 1474–1475 (1971).
30. Hendrickson, W. A. & Lattman, E. E. *Acta crystallogr.* **B26**, 136–143 (1970).
31. Sayre, D. *Acta crystallogr.* **A30**, 180–184 (1974).
32. Hendrickson, W. A. & Konnert, J. H. in *Computing in Crystallography* (eds Diamond, R., Ramaseshan, S. & Venkatesan, K.) 13.01–13.23 (Indian Academy of Sciences, Bangalore, 1980).
33. Konnert, J. H. & Hendrickson, W. A. *Acta crystallogr.* **A36**, 344–350 (1980).
34. Schulz, G. E. & Schirmer, R. H. *Principles of Protein Structure* (Springer, New York, 1979).
35. Templeton, D. H., Templeton, L. K., Phillips, J. C. & Hodgson, K. O. *Acta crystallogr.* **A36**, 436–442 (1980).
36. Phillips, J. C. & Hodgson, K. O. *Acta crystallogr.* **A36**, 856–864 (1980).
37. Karle, J. *Int. J. quant. Chem.* (in the press).
38. Tsernoglou, D. & Petsko, G. A. *FEBS Lett.* **68**, 1–4 (1976).
39. Carter, C. W. Jr, Kraut, J., Freer, S. T. & Alden, R. A. *J. biol. Chem.* **249**, 6339–6346 (1974).
40. Wang, A. H.-J. *et al. Nature* **282**, 680–686 (1979).
41. Karle, I. L., Wieland, T., Schermer, D. & Ottenheim, H. C. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1532–1536 (1979).
42. Kendrew, J. C. *et al. Nature* **185**, 422–427 (1960).
43. Blake, C. C. F. *et al. Nature* **206**, 757–761 (1965).

Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing

Joseph R. Nevins & Michael C. Wilson*

The Rockefeller University, New York, New York 10021, USA

The major late adenovirus-2 transcription unit is active early in infection at a rate equal to that of the other early transcription units. However, transcripts seem to terminate early in infection near map position 60–70 in contrast to late in infection, when termination occurs at map position 99. RNA processing, both poly(A) site selection and splicing, results in the production of a single L1 mRNA during early infection. All these processes are in contrast to those occurring late in infection, indicating that the events of transcriptional termination, poly(A) site selection and splicing can change depending on the conditions of the cell and therefore can participate in the regulation of gene expression.

THE mechanisms that control gene expression in eukaryotic cells are clearly of great importance, but these processes are only poorly understood. The ultimate result of gene regulation, the concentration of functional mRNA in the cytoplasm, could be brought about through the control of transcription, RNA processing and transport, mRNA cytoplasmic stability, or translatability of the mature mRNA. Our aim has been to determine at which of these steps regulation of adenovirus gene expression occurs.

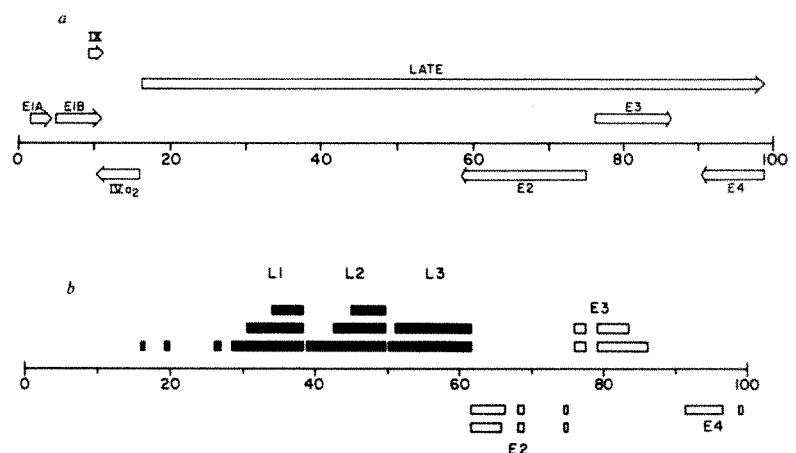
Of the eight known transcription units on the adenovirus chromosome (see Fig. 1), five function primarily early in infection (before DNA replication)^{1–3} while at least two, which are responsible for mRNAs coding for structural proteins IX (refs 3, 4) and IVa2 (ref. 5), become active at intermediate times.

* Present address: Scripps Clinic and Research Foundation, Department of Cellular and Developmental Immunology, La Jolla, California 92037, USA.

The major viral transcription unit active late in infection is responsible for the production of mRNAs coding for most of the structural polypeptides of the virus^{1,6–8}. The initiation site and probably the promoter for this transcription unit, located at map coordinate 16.45 (ref. 9), directs transcription in the rightward direction with termination very near the end of the genome¹⁰. The mRNAs produced from the late transcription unit are arranged into five families of RNAs, L1–L5, the RNAs within a family being 3' co-terminal^{11–14}. Expression of the late transcription unit involves selection of a particular poly(A) site (3' end) among the five possible sites followed by splicing of the tripartite leader sequences to produce one of the mRNAs of the given mRNA family¹⁵.

Although it has long been believed that the expression of the major late transcription unit occurs only after DNA replication, we now show that the late transcription unit is in fact transcribed early in infection but that it is regulated at the point of transcriptional termination in such a way that early in infection only

Fig. 1 *a*, Genomic map locations of the early, intermediate (IVa2 and IX) and late adenovirus transcription units. *b*, Map depicting the genomic positions of the early mRNAs of regions E2, E3 and E4 and the mRNAs of the late 3'-co-terminal families L1, L2 and L3 deriving from the late transcription unit^{11,12,14,17,32}.



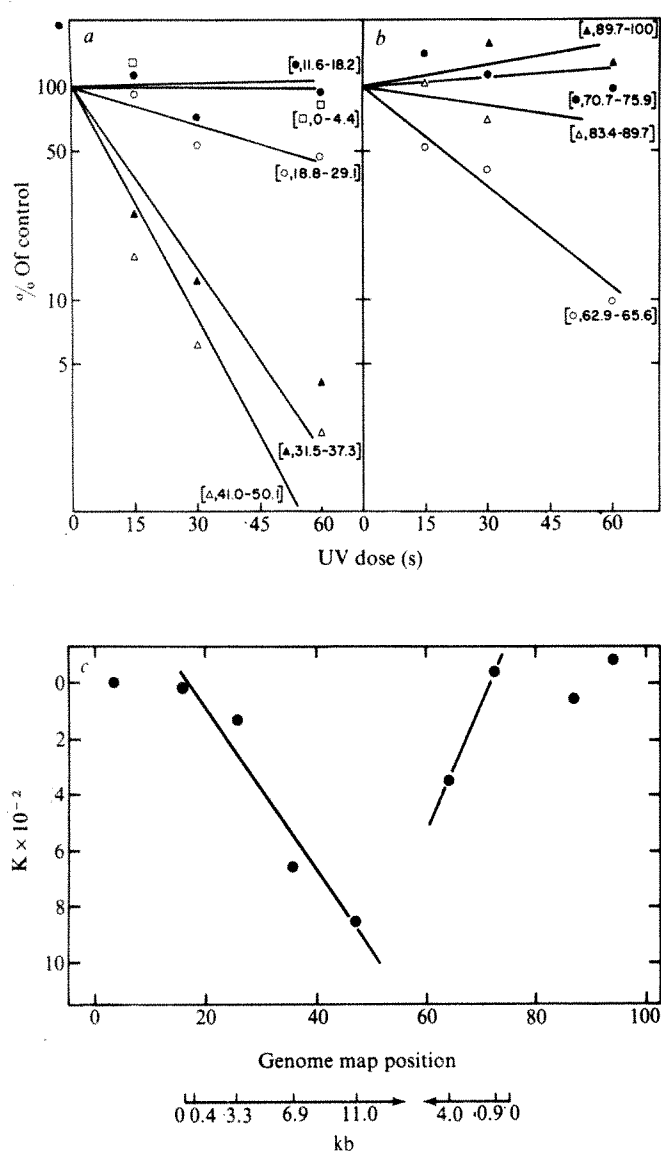


Fig. 2 Sensitivity of early adenovirus RNA synthesis to UV-ray irradiation. *a, b*, At 5 h post-infection, HeLa cells which had been maintained in the presence of $25 \mu\text{g ml}^{-1}$ of cytosine arabinoside were concentrated to 2×10^6 cells per ml. Samples of 2×10^7 cells were UV-ray irradiated for the indicated times¹⁰ and 20 min later were labelled for 5 min with ^3H -uridine. The nuclear RNA was purified and hybridized to filters containing the various DNA fragments prepared as follows: adenovirus-2 DNA restriction fragments of coordinates 0–4.4 (*Hpa* E), 62.9–65.6 (*Bal* K), 70.7–75.9 (*Eco* F), 83.4–89.7 (*Eco* E) and 89.7–100 (*Eco* C) were inserted into the *Pst* site of pBR322 and propagated in *E. coli* HB101. Clones of the *Hind*III I and D fragments were gifts of Dr M. Mathews. A clone of the *Sma* F fragment was a gift of Dr R. Roeder. All procedures for growth of bacteria and handling of recombinant DNAs were in accordance with the NIH Guidelines for Recombinant DNA Research. Plasmid DNAs were purified according to the procedure of Clewell and Hsien³³. Other DNA fragments were prepared by digestion of viral DNA³⁴. In each case, the RNAs were partially broken with alkali before hybridization³⁵. Therefore, any problem arising from possible symmetrical transcription would be eliminated due to denaturation of RNA–RNA duplexes. Such duplexes would then be unlikely to re-anneal due to the vast DNA excess in the hybridizations. Hybridized c.p.m. in the control samples (no UV) were: 454 (0–4.4); 879 (11.6–18.2); 3,330 (18.8–29.1); 1,582 (31.5–37.3); 1,776 (41.0–50.1); 549 (62.9–65.6); 1,534 (70.7–75.9); 351 (83.4–89.7); 747 (89.7–100). *c*, The slopes (*K* value) of the UV inactivation curves were calculated and plotted as a function of genomic map position of the DNA fragment. $K = (\ln R_0/R)/d$ where R_0 = rate of RNA synthesis in the control and R = rates of RNA synthesis at dose d (ref. 19).

the first half is transcribed. We also find that RNA processing, both 3'-end selection (poly(A) site) and 5'-end selection (splicing), is different early and late in infection, which suggests that the regulation of gene expression can in fact be mediated at least in part through RNA processing.

Early transcription of late sequences

Recent reports that sequences corresponding to late RNA can be detected during early infection^{16–18} have thrown little light on their relative frequency or the extent of transcription. Accordingly, we have used the technique of UV-ray irradiation transcription mapping to investigate the early transcription of various regions of the adenovirus-2 genome. The principle is that the transcriptional effectiveness of a sequence in an irradiated DNA template decreases exponentially with increasing distance from the promoter¹⁹. The transcriptional activity of the various parts of the genome can be measured by hybridizing the RNA labelled during 5-min pulses of ^3H -uridine with DNA fragments corresponding to defined sections of the genome.

HeLa cells early in infection (5 h post-infection) in the presence of cytosine arabinoside to prevent viral DNA replication) were treated with various doses of UV rays and then with ^3H -uridine. The labelled nuclear RNA was hybridized to a variety of DNA fragments from the adenovirus genome (Fig. 2*a, b*). Figure 2*a* presents the results for sequences between 11 and 50 map units and Fig. 2*b* those for sequences to the right of 62 map units. The data obtained were plotted to compare the sensitivity to UV-ray inactivation of the various sequences (the slopes of the UV-ray inactivation curves) as a function of their position in the genome (Fig. 2*c*). Region E2 is represented by the DNA fragments at 62.6–65.9 map units, and 70.7–75.9 map

Table 1 Extent of rightward transcription during early infection

Expt	DNA	Map coordinates	Hybridized c.p.m.	c.p.m. per kilobases of hybrid
1*	<i>Sma</i> I	36.7–40.5	883	645
	<i>Kpn</i> I	39.6–42.7	495	442
	<i>Kpn</i> H	42.7–47.4	1,555	920
	<i>Sma</i> H	52.6–56.9	1,120	723
2*	<i>Sma</i> B/ <i>Bam</i> B	18.8–29.1	982	265
	<i>Sma</i> A/ <i>Eco</i> A (rightward)	56.9–58.5	149	266
	<i>Sma</i> A/ <i>Eco</i> A (leftward)	56.9–58.5	182	325
	<i>Eco</i> F	70.7–75.9	850	538
3†	<i>Sma</i> I	36.7–40.5	405	337
	<i>Sma</i> H	52.6–56.9	553	357
	<i>Bal</i> K (rightward)	62.9–65.6	69	138
	<i>Bal</i> K (leftward)	62.9–65.6	169	338
	<i>Eco</i> F	70.7–75.9	489	309

At the appropriate time post-infection cells were concentrated to 1×10^7 cells per ml in fresh, warm minimal essential medium. ^3H -uridine was added at $200 \mu\text{Ci ml}^{-1}$ and then incorporation of label was halted by pouring the cells on to crushed frozen phosphate-buffered saline. Nuclear RNAs were then extracted as described previously^{29–31}. The nuclear RNA was hybridized to the various DNA fragments immobilized on nitrocellulose filters¹². The source of the DNAs were described in Fig. 2 legend. To determine relative rates of transcription, each value for hybridized c.p.m. was corrected for the size of the hybrid (the size of the DNA fragment). One map unit was assumed to represent 350 nucleotides. The exception is the *Bal* K fragment, which represents 2.7 map units or 0.97 kilobases. The DNA used in this work was a clone of *Bal* K in which only 0.5 kilobases of the fragment was obtained in the cloned plasmid.

* At 5 h post-infection, a sample of 2×10^7 cells which had been maintained in cytosine arabinoside ($25 \mu\text{g ml}^{-1}$) was concentrated to 1×10^7 cells per ml and labelled with ^3H -uridine ($200 \mu\text{Ci ml}^{-1}$) for 5 min. Nuclear RNA was prepared and hybridized to the various DNA fragments.

† Same as in *a* except that cells were labelled at 3.5 h post-infection.

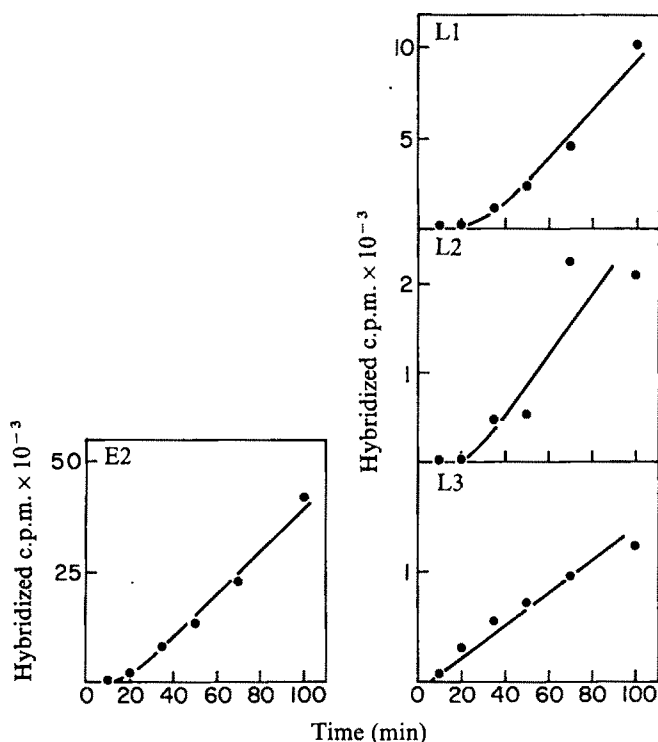


Fig. 3 Kinetics of accumulation of ^3H -uridine into cytoplasmic poly(A) $^+$ RNA hybridizing to early region 2 DNA (left panel) and to L1, L2 and L3 DNAs (right panel). At 5 h post-infection, cells which had been maintained in the presence of $25 \mu\text{g ml}^{-1}$ of cycloheximide were concentrated to 4×10^6 cells per ml. ^3H -uridine was added ($200 \mu\text{Ci ml}^{-1}$) and samples of 5×10^7 cells were taken at the indicated times. Cytoplasmic RNA was purified from each sample²⁹⁻³¹, selected on poly(U)-Sephacrose³⁶ and hybridized to the appropriate DNA fragments immobilized on nitrocellulose filters¹². E2 (*Bal K*, 62.9–65.6), L1 (*HindIII I*, 31.5–37.3), L2 (*HindIII D*, 41.0–50.1), L3 (*Sma H*, 52.6–56.9).

units with the curve originating at the region of greatest resistance to UV-ray irradiation, the promoter at coordinate 75. Transcription of the other regions of the genome containing early transcription units (regions 1A, 3 and 4) was also found to be very resistant to UV-ray inactivation as expected considering the proximity of these particular sequences to their promoters. The transcription detected between map positions 16 and 50 apparently derives from the region of map position 16. Note that the slope of this curve is opposite to that of the early region 2 curve, indicating transcription proceeding in the opposite direction. These results therefore indicate that the region between 16 and 50 map units is transcribed as a single unit, reading in the rightward direction and originating near map position 16. Fingerprint analysis of RNA made early in infection and specific to the 16–18 region indicated that the major late promoter was indeed active²⁰. Furthermore, the RNA detected by electron microscopy during early infection both in the cytoplasm¹⁶ and in the nucleus¹⁷ possessed a 5' terminus which mapped to genome coordinate 16. It is therefore likely that the same promoter at map position 16.45 which directs transcription during late infection is responsible for the synthesis of RNA which we have detected during early infection. In addition, comparison of the hybridized c.p.m. in the control samples (no UV) of Fig. 2 for the various regions of the genome indicates that the late sequences are transcribed at a rate equal to that of any of the early transcription units, a result also obtained by Shaw and Ziff²⁰.

As shown in the experiment in Fig. 2, transcription in the rightward direction originating at the late promoter at map position 16 does not extend to map position 70 and beyond. For example, transcription of the 83.4–89.7 region originating at map unit 16.45 would have a UV target size of 25 kilobases,

yielding a K value in Fig. 2 of ~ 21 . Rather, the value of 0.6 which was obtained suggests a very small UV target size, fully consistent with the hypothesis that all this transcription is derived from the early region 3 promoter. This argument can be made because in this experiment the rate of transcription of the late sequences between map coordinates 18 and 50 was approximately four times that of the 83.4–89.7 region. Thus, even a very small amount of readthrough from the late region into the region of 83.4–89.7 would have greatly increased the UV-ray sensitivity. In addition, transcription in the rightward direction of the sequences between 70.7 and 75.9 map units originating at or near the promoter at coordinate 16 would have a UV target size of 20 kilobases, yielding a K value in Fig. 2 of ~ 16 ; thus, as transcription of the late region was proceeding at a rate equal to that of early region 2 in this particular experiment, transcription of 70.7–75.9 DNA sequences in both the rightward and leftward directions would yield an average K value of close to 8, which clearly was not the case. This suggests that transcription in the rightward direction terminated somewhere between 50.1 and 70.7 map units before reaching the 70–75 region. To define this site more accurately, pulse-labelled nuclear RNA was hybridized to separated strands of the *Eco A/Sma A* fragment (map coordinates 56.9–58.5) and the *Bal K* fragment (map coordinates 62.9–65.6, Table 1). The leftward-reading strand of each fragment detects early region 2 transcription²¹; transcription of each of these was approximately equimolar to that of the *Eco F* fragment (70.7–75.9). The rightward-reading strand of the *Sma A/Eco A* fragment also hybridized pulse-labelled nuclear RNA; the hybridized radioactivity, when compared with that for fragments to the left, indicated transcription to be equimolar from 16.45 through the 56.9–58.5 region. In a separate experiment, the rightward-reading strand of the *Bal K* fragment also hybridized labelled nuclear RNA but only about 40% of that which hybridized to the leftward-reading strand of the *Bal K* fragment. Because transcription of the *Sma I* (36.7–40.5), *Sma H* (52.6–56.9), *Eco F* (70.7–75.9) and leftward *Bal K* (62.9–65.6) fragments all appeared to be equimolar, it is possible that transcription in the rightward direction terminates

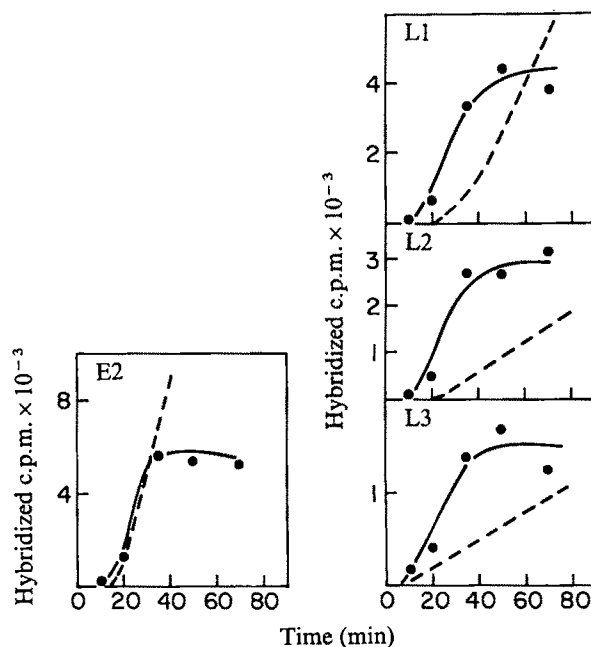


Fig. 4 Kinetics of accumulation of ^3H -uridine into nuclear poly(A) $^+$ RNA hybridizing to early region 2 DNA (left panel) and to L1, L2 and L3 DNAs (right panel). Nuclear RNA was prepared from the experiment described in Fig. 3 legend, selected on poly(U)-Sephacrose and hybridized to the same DNA probes described in Fig. 3 legend. The cytoplasmic RNA accumulation curves (---) from Fig. 3 have been plotted for comparison.

Table 2 Rates of accumulation of cytoplasmic and nuclear poly(A)⁺ RNAs

DNA region	Cytoplasmic rate (c.p.m. per kilobase per min)	Nuclear rate (c.p.m. per kilobase per min)	Fraction transported*
E2	980	800	1.23
L1	69	75	0.92
L2	10	39	0.26
L3	10	32	0.31

Cytoplasmic rate was determined from the data presented in Fig. 3, and the nuclear rate from the data presented in Fig. 4, in each case using the linear portion of the curves.

* Cytoplasmic rate/nuclear rate.

within the *Bal K* fragment if, in fact, a single termination site exists. However, it is clear that transcription termination prevents expression of at least the final 10 kilobases (map units 70–99) of the late transcription unit during early infection. This is in sharp contrast to the situation late in infection when termination occurs near the end of the genome^{10,15}.

Formation of cytoplasmic mRNA

We measured the rate of accumulation of the first three mRNA families which are included in the first half of the major late transcription unit (see Fig. 1). Infected cells maintained in the presence of cycloheximide were labelled continuously with ³H-uridine from 5 h post-infection. Samples were taken after various times of labelling, and cytoplasmic poly(A)⁺ RNA was prepared and hybridized to DNA probes representing the first three late mRNA families and to a probe for early region 2 as a comparison (Fig. 3). RNA specific to each of these late mRNA families was found accumulating in the cytoplasm although more slowly than that from early region 2. The rates of accumulation were determined for each of the regions using the data presented in Fig. 3 normalized to take into account the sizes of the various DNA probes (Table 2). The L1 family accounted for ~80% of the late RNA accumulating in the cytoplasm during early infection and was about 7% of the rate of the early region 2 mRNA. However, in this particular experiment the rate of transcription of the E2 sequences, as measured in a 10-min ³H-uridine labelling test, was 2.8 times that of the L1 sequences. Thus, after correcting the cytoplasmic accumulation rates for this transcription discrepancy, the L1 mRNA accumulation was 20% the rate of the early region 2. However, even with the correction for relative transcription rates, L2 and L3 mRNA production was well below that of early region 2 (~3%).

Processing of the nuclear RNAs

The preponderance of L1 RNA sequences in the cytoplasm during early infection (a result also seen by others using both electron microscopy¹⁶ or pulse-labelling²⁰) despite equimolar transcription of each of the first three families (see Table 1) implies a post-transcriptional regulation event: either a bias in poly(A) site selection favouring the L1 site or a failure to transport polyadenylated L2 and L3 RNAs to the cytoplasm. This question was analysed by comparing the rates of formation of nuclear and cytoplasmic poly(A)-containing RNAs. The experiment shown in Fig. 3 was also the source for nuclear poly(A)⁺ RNA which was hybridized to DNA probes for the L1, L2 and L3 regions and to a probe for early region 2 (Fig. 4). Clearly, as indicated by the equivalent slopes of the two curves, the early region 2 nuclear poly(A)⁺ RNA was fully converted into cytoplasmic RNA. That is, the rate of accumulation of early region 2 poly(A)-containing nuclear RNA is paralleled closely

in time by an equal rate of accumulation of the sequences in the cytoplasm. An inspection of the accumulation curves of nuclear and cytoplasmic poly(A)-containing RNAs specific to regions L1, L2 and L3 suggested a nearly total conservation of the L1

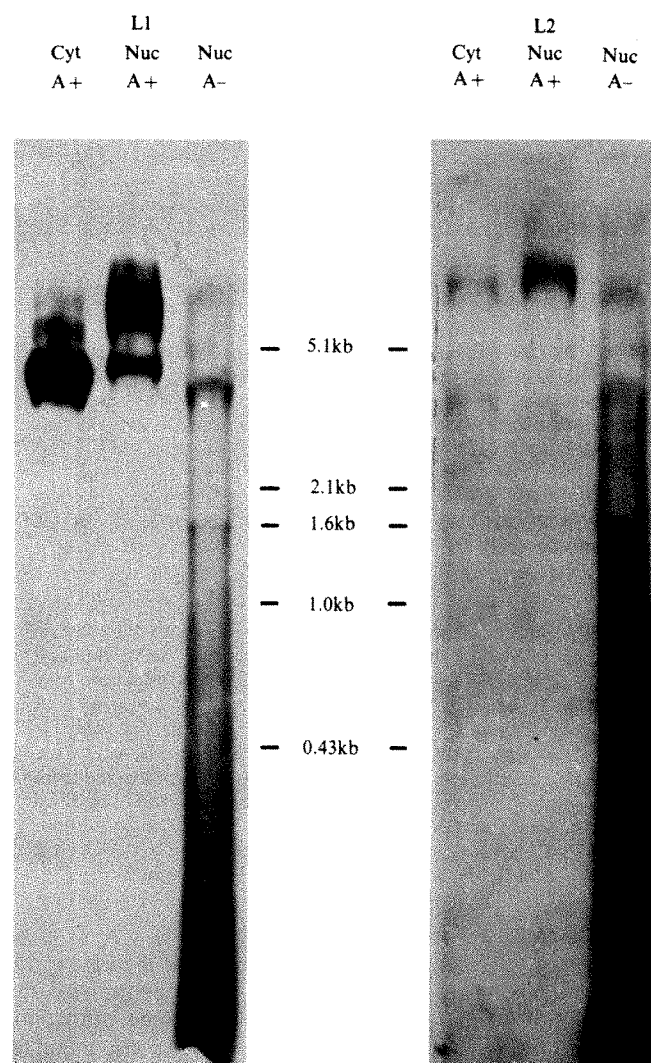
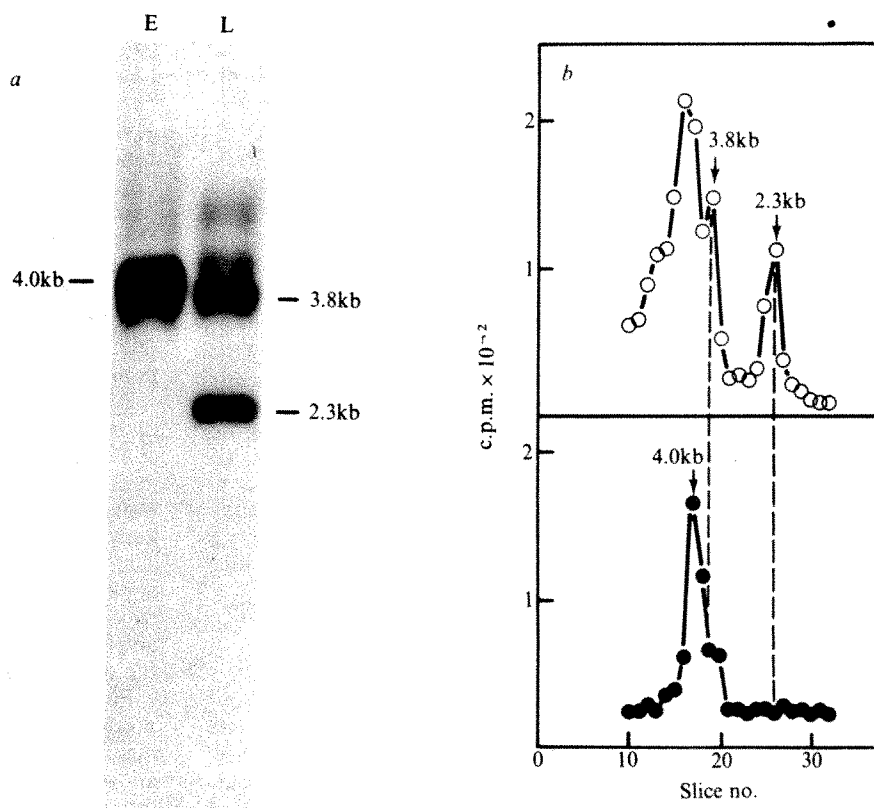


Fig. 5 Nuclear and cytoplasmic RNAs from early infection detected with L1- and L2-specific probes. Nuclear and cytoplasmic RNA was prepared from cells at 5 h post-infection which had been maintained in cycloheximide (25 µg ml⁻¹). Poly(A)⁺ nuclear and cytoplasmic RNAs and the poly(A)⁻ nuclear RNA were denatured by heating to 65 °C for 5 min in 50% formamide, 1.1 M formaldehyde and 1× RB (20 mM MOPS pH 7.0, 5 mM Na acetate, 1 mM EDTA). After addition of bromophenol blue and glycerol, the samples were loaded on to a 1.5% agarose gel containing 1× RB and 1.1 M formaldehyde. Electrophoresis was at 50 V for about 10 h. The gel was soaked in 20× SSC for 60 min and then blotted to a nitrocellulose sheet (Schleicher & Schuell BA85) according to the method of Southern^{37,38}. The RNAs were detected by hybridization with ³²P *Hind*III IDNA (L1) or *Hind*III D DNA (L2). Plasmid DNAs were labelled by nick-translation to a specific activity of 0.5–1.0 × 10⁸ c.p.m. per µg essentially by the procedure of Rigby *et al.*³⁹. The nitrocellulose sheets were prehybridized overnight in 0.75 M NaCl, 50 mM TESS pH 7.0, 50 mM EDTA, 50% formamide, 5× Denhardt's solution⁴⁰ and 100 µg ml⁻¹ denatured salmon sperm DNA at 45 °C. Hybridization was carried out in the same conditions for 15–20 h but with 1× Denhardt's solution, 100 µg ml⁻¹ salmon sperm DNA and ~1–5 × 10⁶ c.p.m. ml⁻¹ ³²P DNA. After hybridization, the filters were washed successively with hybridization buffer at 45 °C and a final wash with 50% formamide, 10 mM Tris pH 7.4 and 10 mM EDTA at 45 °C for 30 min. The sheets were exposed to preflashed Cronex X-ray film with an intensifying screen for 20–60 h at –70 °C. kb, Kilobases.

Fig. 6 Comparison of L1-specific RNA species produced early and late in infection. *a*, Poly(A)⁺ cytoplasmic RNA was prepared from cells at 5 h post-infection (in the presence of cycloheximide) (E) and 15 h post-infection in the absence of drug (L). The samples were denatured, fractionated in a 1.2% agarose-formaldehyde gel, blotted to nitrocellulose and hybridized with ³²P-labelled *Hind*III I DNA (L1-specific) as described in Fig. 5 legend. *b*, Cells at 5 h post-infection (in the presence of cycloheximide) were labelled with ³H-uridine (200 μ Ci ml⁻¹) for 45 min; cells at 15 h post-infection were labelled with ³²P for 60 min. Poly(A)⁺ cytoplasmic RNA was prepared from each, mixed together and fractionated in a 2.2% acrylamide-0.5% agarose gel³². After electrophoresis, the gel was divided into 2-mm slices, RNA was eluted and then hybridized to filters bearing the *Sma* I fragment (map coordinates 36.7–40.5). Upper panel (○), ³²P late mRNA; lower panel (●), ³H early mRNA. kb, Kilobases.



RNA but a less than complete nuclear to cytoplasmic transport of the L2 and L3 species. Once again, the data were used to determine rates of accumulation for comparison (Table 2). Whereas the L1 sequences accumulated in the cytoplasm at a six- to sevenfold higher rate than for the L2 and L3 sequences, the rates of accumulation in the nucleus for the three regions were within a factor of about two. Thus, the higher rate of appearance of L1 mRNA sequences in the cytoplasm is due to two factors: a bias in the selection of the L1 poly(A) site and a failure to transport all polyadenylated L2 and L3 sequences.

Differential splicing of L1 sequences during early and late infection

Nuclear and cytoplasmic RNA was prepared from cells at 5 h post-infection (in the presence of cycloheximide from the time of infection). The nuclear poly(A)⁺ and poly(A)⁻ RNA and cytoplasmic poly(A)⁺ RNA were electrophoresed in an agarose-formaldehyde gel, transferred to nitrocellulose and analysed by hybridization to a ³²P nick-translated DNA probe specific for the L1 region (Fig. 5). For comparison, L2-specific RNA was also hybridized with a ³²P-labelled *Hind* D probe. Both the probes detected nuclear RNA; the poly(A)⁻ RNAs detected with each probe were mostly less than about 500 nucleotides long, and were presumably breakdown products. Discrete poly(A)⁺ nuclear RNAs were detected. In addition, a discrete species of poly(A)⁺ cytoplasmic RNA, 4.0 kilobases long, was detected with the L1 probe and two species of cytoplasmic RNA, 5.8 and 3.3 kilobases, were detected with the L2-specific probe. However, the L2 cytoplasmic RNA was present in very low abundance compared with the L1-specific RNA, in agreement with the data presented in Fig. 3 and Table 2. The cytoplasmic L1-specific poly(A)⁺ RNA also appeared to be more abundant at steady state than the nuclear poly(A)⁺ L1 RNA, whereas the reverse was true for the L2-specific nuclear and cytoplasmic RNAs. This is consistent with the kinetic analysis result that not all the L2 poly(A)-containing nuclear

RNA is transported to the cytoplasm (see Fig. 4, Table 2). Because of the very low abundance of the L2 RNAs we have not characterized them further.

The finding that only one species of RNA seemed to be formed from the L1 region during early infection was of interest because it contrasted with the situation late in infection where multiple RNAs are formed from this region^{11,12,14}. For a direct comparison, poly(A)⁺ mRNA was prepared from late infected cells and from cells at 5 h post-infection in the presence of cycloheximide. These two samples were then fractionated in an agarose-formaldehyde gel, transferred to nitrocellulose and hybridized with ³²P-*Hind*III I DNA (map coordinates 31.5–37.3) (Fig. 6a). Two major species of late RNA (3.8 and 2.3 kilobases) and two minor species (5.6 and 4.3 kilobases) were detected with the L1 probe. It is possible that the larger of the minor species represents leakage of a precursor molecule from the nucleus. Clearly, however, only one species was detected in the RNA prepared from early infected cells. This RNA was again 4.0 kilobases long and thus did not correspond to any of the RNAs from the late preparation although it may represent the 3.8-kilobase late RNA with an additional leader segment derived from map position 22 as reported by Chow *et al.*¹⁶. Regardless of the actual structure of this RNA, it is clear that the 2.3-kilobase species formed at late times of infection is not produced from the same sequences of DNA early in infection.

However, because this experiment examined only the steady-state population of RNA, we could not say whether the difference was due to differential splicing or differential stability. That is, it was possible that the 2.3-kilobase L1 mRNA was actually formed early in infection but was either not transported to the cytoplasm or was rapidly degraded on reaching the cytoplasm. A failure to transport a properly spliced 2.3-kilobase RNA seems not to be the case because we could not detect this species in the nuclear poly(A)⁺ RNA (Fig. 5) whereas late in infection the 2.3-kilobase RNA could be detected in the nucleus (data not shown). The possibility of rapid cytoplasmic turnover of 2.3-kilobase RNA seemed unlikely as the accumulation data of Fig. 3 gave no indication of instability of the total L1

sequences during 100 min of labelling. Rapid turnover of one half of the L1 mRNA (the 2.3-kilobase species) should have been reflected in the cytoplasmic accumulation curve. However, to answer this question directly, RNA was labelled with ^3H -uridine for 40 min at 5 h post-infection. The poly(A)⁺ cytoplasmic RNA was mixed with a sample of ^{32}P late RNA, fractionated in an acrylamide-agarose gel and hybridized with an L1-specific probe (Fig. 6b). Once again, it was clear that both a 3.8- and a 2.3-kilobase species were formed late in infection. The additional higher molecular weight RNA species detected in this gel represent the minor L1 RNA (Fig. 6a) and the largest RNA of the L2 group because the probe used for this hybridization extends to map position 40.5 (see ref. 12). The pulse-labelled RNA from early-infected cells, however, only contained the 4.0-kilobase species, in agreement with the analysis of steady-state RNA. The clear result both from the pulse-labelled sample and from the steady-state preparations is the absence of the 2.3-kilobase L1 RNA during early infection. Therefore, we conclude that the reason for the qualitative change in L1 RNAs during early and late infection is probably a change in processing at the level of RNA splicing.

Discussion

These results allow us to draw several new conclusions concerning adenovirus gene expression and its regulation. First, the late transcription unit is not under negative control during early infection. Although the rate of transcription of the late transcription unit during early infection is much less than that during late infection, it is still equal to the rate of any of the other early transcription units. Termination of transcription, rather than initiation, may be considered to be a controlling factor in the expression of the late transcription unit during the early phase of infection. Before DNA replication the late transcription unit seems to be active but only to produce RNA corresponding to sequences from the L1, L2 and L3 late mRNA families. Termination of transcription prevents the expression of the remainder of the transcription unit; the mechanism for this event early in infection is not clear although it is noteworthy that transcription proceeds from the late promoter approximately as far as the early region 2 transcription unit. It is possible that the opposing transcription of early region 2 prevents further transcription in the rightward direction and results in termination. This block would then be alleviated late in infection after the synthesis of new viral templates and when early region 2 transcription had been shut off^{22,23}. Alternatively, the termination of transcription in the middle of the genome during early infection or the prevention of this termination late in infection (antitermination) could be an event mediated by a specific factor.

Second, processing in the nucleus of the primary transcripts from the late sequences favours the L1 mRNA region. This is in contrast to the situation during late infection, when the L1

sequences are produced in lowest amounts of the five groups of late mRNAs¹⁵. For instance, during early infection the ratio of L1 to L2 sequence selection in the nucleus is 2:1; however, during late infection the ratio is 1:2. Thus, the process which governs the selection of 3' ends seems to vary according to the conditions of the cell. This finding is also significant in relation to the mechanism by which the poly(A) sites are selected. The finding that the selection of a poly(A) site seems to be changeable implies that the specificity does not reside in the DNA sequence, but rather, that a factor, specific for a given poly(A) site, is responsible for the selection. A change in the concentration of the factor, as a function of time of infection, would then result in changes in the frequency of selection of the possible sites.

Are the processes involved in regulating poly(A) site selection of any importance to the eukaryotic cell? Clearly, such an event could only be valuable to those transcriptional units possessing multiple poly(A) sites. It has recently been shown that the transcription unit for the secreted and membrane-bound IgM heavy chains contains two separate poly(A) addition sites^{24,25}. The basis for selection between those two gene products apparently resides at the choice of the 3' end of the mRNA and could be regulated in their expression by either a termination-anti-termination mechanism or by a changing selection of one of the two sites, as discussed above.

Finally, our results concerning the species of L1 RNAs produced early and late in infection demonstrate that processing, apparently at the level of RNA splicing, can be altered to allow different RNAs to be produced from the same DNA sequence. These results thus indicate that the splicing mechanism changes depending on the environment of the cell (in this case early or late in infection) and therefore can participate in the regulation of gene expression. The nature of the change causing differential splicing is unknown, but clearly involves some difference between an early-infected and a late-infected cell. Recently, there has been much interest in the possible role of the small nuclear RNAs in the splicing of cellular hnRNA^{26,27}. The adenovirus VA RNAs have been postulated to perform a role in splicing of the late viral RNAs in a manner analogous to that proposed for the cellular small nuclear RNA²⁸. If this were true, it may be that early in infection the L1 RNAs are processed utilizing the cellular small nuclear RNA rather than VA RNA. Once VA RNA is formed, the pattern then changes to the set of RNAs observed late in infection. Of course, it is equally possible that a viral protein is required for the splicing reaction which produces the late set of L1 mRNAs. Regardless of the exact nature of the mechanism, it seems clear that a system is available to probe such a question.

We thank Dr J. E. Darnell for helpful discussions, Dr R. Roeder for the clone of the *Sma* F fragment and Dr M. Mathews for the clones of the *Hind*III I and D fragments. We also thank Janet Winkler for technical assistance. This work was supported by NIH grant GM 26765-02. J.R.N. is the recipient of NCI Research Career Development Award CA00666-01.

Received 8 September 1980; accepted 13 January 1981.

1. Evans, R. M. *et al.* *Cell* **12**, 733-739 (1977).
2. Berk, A. J. & Sharp, P. A. *Cell* **12**, 45-55 (1977).
3. Wilson, M. C., Fraser, N. W. & Darnell, J. E. *Virology* **94**, 175-184 (1979).
4. Alestrom, P. *et al.* *Cell* **19**, 671-681 (1980).
5. Chow, L. T. & Broker, T. *Cell* **15**, 497-510 (1978).
6. Goldberg, S., Weber, J. & Darnell, J. E. *Cell* **10**, 617-621 (1977).
7. Weber, J., Jelinek, W. & Darnell, J. E. *Cell* **10**, 612-617 (1977).
8. Goldberg, S., Nevins, J. & Darnell, J. E. *J. Virol.* **25**, 806-810 (1978).
9. Ziff, E. & Evans, R. *Cell* **15**, 1463-1475 (1978).
10. Fraser, N. W., Nevins, J. R., Ziff, E. & Darnell, J. E. *J. molec. Biol.* **129**, 643-656 (1979).
11. McGrogan, M. & Raskas, H. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 625-629 (1978).
12. Nevins, J. R. & Darnell, J. E. *J. Virol.* **25**, 811-823 (1978).
13. Ziff, E. & Fraser, N. W. *J. Virol.* **25**, 897-906 (1978).
14. Chow, L. T., Roberts, J. M., Lewis, J. B. & Broker, T. R. *Cell* **11**, 819-836 (1977).
15. Nevins, J. R. & Darnell, J. E. *Cell* **15**, 1477-1493 (1978).
16. Chow, L. T., Broker, T. R. & Lewis, J. B. *J. molec. Biol.* **134**, 265-303 (1979).
17. Kitchingham, G. R. & Westphal, H. J. *J. molec. Biol.* **137**, 23-48 (1980).
18. Lewis, J. B. & Mathews, M. B. *Cell* **21**, 303-313 (1980).
19. Sauerbier, W. *Adv. Radiat. Biol.* **6**, 49-106 (1976).
20. Shaw, A. R. & Ziff, E. B. *Cell* **22**, 905-916 (1980).
21. Nevins, J. R., Blanchard, J.-M. & Darnell, J. E. *J. molec. Biol.* **144**, 377-386 (1980).

22. Nevins, J. R., Ginsberg, H. S., Blanchard, J.-M., Wilson, M. C. & Darnell, J. E. *J. Virol.* **32**, 727-733 (1979).
23. Nevins, J. R. & Jensen-Winkler, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1893-1897 (1980).
24. Alt, F. W. *et al.* *Cell* **20**, 293-301 (1980).
25. Rogers, J. *et al.* *Cell* **20**, 303-312 (1980).
26. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283**, 220 (1980).
27. Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1877-1879 (1980).
28. Mathews, M. *Nature* **285**, 575-577 (1980).
29. Penman, S. *J. molec. Biol.* **17**, 117-130 (1966).
30. Penman, S., Scherrer, K., Becker, Y. & Darnell, J. E. *Proc. natn. Acad. Sci. U.S.A.* **49**, 654-662 (1963).
31. Soeiro, R. & Darnell, J. E. *J. molec. Biol.* **44**, 551-562 (1969).
32. Berk, A. J. & Sharp, P. A. *Cell* **14**, 695-711 (1978).
33. Clewell, D. B. & Helinski, D. R. *Biochemistry* **9**, 4428-4440 (1970).
34. Nevins, J. R. *Meth. Enzym.* **65**, 768-785 (1980).
35. Jelinek, W., Molloy, G., Fernandez-Munoz, R., Saiditt, M. & Darnell, J. E. *J. molec. Biol.* **82**, 361-370 (1974).
36. Molloy, G. G., Jelinek, W., Saiditt, M. & Darnell, J. E. *Cell* **1**, 43-53 (1974).
37. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
38. Derman, E. *et al.* *Cell* (in the press).
39. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237-251 (1977).
40. Denhardt, D. *Biochem. biophys. Res. Commun.* **23**, 641-652 (1966).

LETTERS

UV observations of X-ray sources
2A0311-227 and 2A0526-328

M. J. Coe

Physics Department, Southampton University,
Southampton SO9 5NH, UK

D. T. Wickramasinghe

Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

The AM Herculis-type of dwarf novae have emerged as a new class of cataclysmic variables. They are characterized by fairly steady accretion from a late-type primary on to a strongly magnetized white dwarf. The progenitor of this group, AM Herculis, shows strong optical polarization¹ and prominent emission lines in both the optical² and the UV³. It is also a hard X-ray⁴⁻⁶ and an IR source². Two other X-ray sources, 2A0311-227 (ref. 7) and 2A0526-328 (ref. 8), have recently been suggested as members of the AM Her class on the strength of their similar optical spectra. To compare these objects with AM Her and to extend investigations of this new class of dwarf novae the International Ultraviolet Explorer (IUE) satellite was used to obtain their UV spectra and to search for any changes with the binary motion. We show here that the results confirm the identification of 2A0311-227 with the AM Her class, but not 2A0526-328 which seems to exhibit a UV spectrum more consistent with an SS Cygni type object. Changes in the emission lines' intensities of at least a factor of two were observed as a function of the binary phase.

The IUE observatory⁹ has long wavelength (1,900-3,200 Å) and short wavelength (1,100-1,900 Å) cameras and two resolution modes (10 and 0.2 Å). The images obtained in this observation were taken during an 8-h period starting at 20.00 UT on 24 July 1980 and were all at low resolution and with the large aperture (Table 1). At the time of the observations 2A0526-328 was fainter than $B = 14.5$, while 2A0311-227 was at $B \approx 14.2$. The emission lines were well exposed on both sources and the continua satisfactorily so at least in the case of 2A0526-328. There was, however, some contamination of the long wavelength spectrum of 2A0526-328 at $\lambda > 2,900$ Å probably arising from scattered Earth light.

Figure 1a shows one of the two short wavelength spectra of 2A0311-227. This source has a binary period of 81 min and the data in Fig. 1a were integrated over binary phases 0.6 to 0.1 with phase 0.0 corresponding to maximum light¹⁰. The more prominent line features have been identified and are seen sitting on a fairly flat continuum. This spectrum is very similar to that seen³ from AM Her with all of the spectral lines being common. In general, the UV emission is 5-10 times weaker than in AM Her, which could be partially due to different source distances. The second short wavelength spectrum taken about an hour later over binary phases 0.1 to 0.6 shows the emission lines to be about 40% brighter during this half of the binary cycle. There is no evident change in the level of the continuum, but these images are not well-enough exposed to be too positive about this aspect. The overall shape of the continuum is consistent with a Rayleigh-Jeans $F_{\nu} \propto \nu^{-2}$ law similar to that found by Raymond *et al.* from AM Her. The long wavelength images are almost featureless with no evidence for the interstellar feature at 2,200 Å, indicating that the source distance is ≤ 500 pc.

In contrast, the short wavelength spectrum of 2A0526-328 (Fig. 1b) shows a considerably brighter UV continuum although it is fainter than 2A0311-227 in the optical. Although the emission lines are far less prominent most of the lines visible in

2A0311-227 can also be identified in this source. The significant difference is the presence of absorption or P Cygni profiles seen just resolved in the C IV, S IV and N V resonance lines. These have not been observed in AM Her or 2A0311-227, but some evidence for their existence was found in the UV spectrum of another dwarf nova, SS Cygni, while in an active state¹¹. It is not believed, however, that 2A056-328 was in a particularly bright state during these UV observations because the fine error sensor on board IUE, which has a limiting magnitude⁹ of $B \approx 14.5$, did not detect the source. Charles *et al.*⁸ give its brightness as $V = 13.5$ on 26 October 1978, significantly more than during the current observation. If these absorption features are interpreted in terms of a gas outflow from the system, then an average velocity may be measured for such motion from all the observed features assuming that the absorption and emission components are just resolved. This is found to be $2,000 \pm 600$ km s⁻¹. A second short wavelength spectrum taken about an hour later (the binary period is still uncertain, but is probably in the range 4-5 h (ref. 12)) shows the emission features considerably reduced with only C IV still obvious, but the absorption dips remain essentially unchanged.

The UV continuum of 2A0526-328 is well-fitted by a black body with temperature of $(2.25 \pm 0.1) \times 10^4$ K. This value is similar to that seen by Heap *et al.*¹¹ in SS Cygni who deduced a temperature of 3×10^4 K for this source in its quiescent state. It is also very similar to the black-body temperature of 2.4×10^4 K found by Bath *et al.*¹³ in another X-ray cataclysmic variable, EX

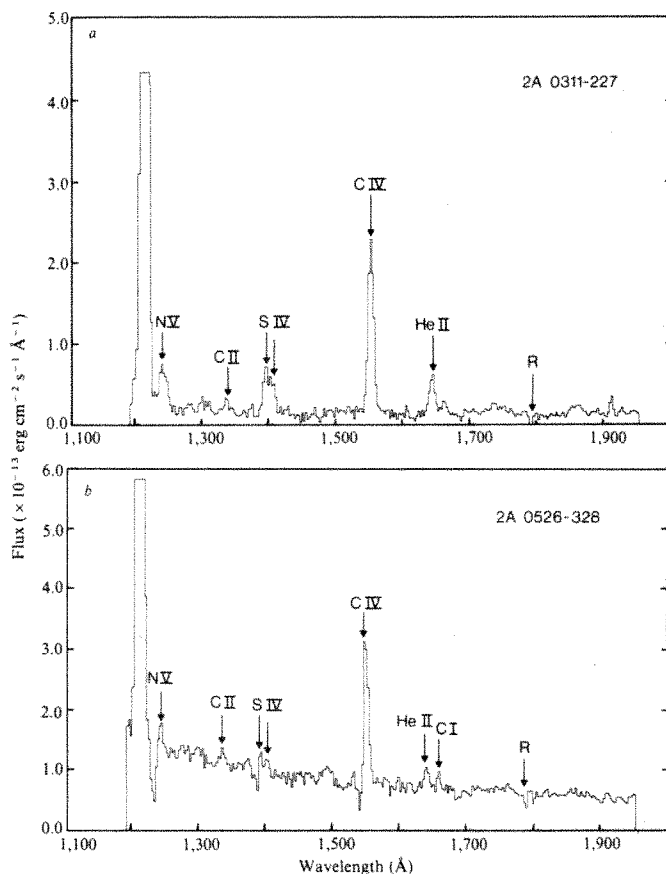


Fig. 1 a, The short wavelength UV spectrum of 2A0311-227 taken by IUE on 25 July 1980. The exposure was 40 min and started at 00.42 UT. b, The short wavelength UV spectrum of 2A0526-328 taken by IUE on 24 July 1980. The exposure was 30 min starting at 21.12 UT. R is a reseau or calibration mark and the large feature at 1,216 Å is due to geocoronal H Ly α emission.

Table 1 Observations carried out during the 8-h IUE shift starting at 20.00 UT on 24 July 1980

Source	Start time (UT)	Exposure (min)	Wavelength	Image no.
2A0526-328	21.12	30	Short	39590
	21.46	30	Long	28336
	22.13	30	Short	39591
2A0311-227	23.40	30	Long	28337
	00.42	40	Short	39592
	01.31	50	Long	28338
	02.44	40	Short	39593

Hydrae. Fitting the $F_{\nu}\nu^2$ law to the continuum of 2A0526-328 is totally unacceptable. However, a $F_{\nu}\nu^{0.3}$ energy distribution predicted from optically thick black-body disk models is generally consistent with these UV observations and the available lower limits on the B magnitude ($B \geq 14.5$). This is another significant difference from the AM Her-type sources where the UV and optical emission seems to arise from different components.

Thus, while 2A0311-227 and 2A0526-328 are both cataclysmic variables their UV spectra show several distinct

differences. The source 2A0311-227 has many similarities with AM Her and is probably a member of this class. Its UV continuum is consistent with $F_{\nu}\nu^2$ and as such probably represents the Raleigh-Jeans tail of a black body peaking in very soft X rays ($E \sim 0.1$ keV). This is similar to the conclusion reached by Raymond *et al.*³ and Fabbiano *et al.*¹⁴ for AM Her. Fabbiano *et al.* conclude that in the case of AM Her the strong magnetic field will give rise to an accretion column as the source of the UV/soft X-ray emission. This will almost certainly also be the case with 2A0311-227 for which one can deduce a comparable magnetic field from its strong optical polarization¹⁵. In the case of 2A0526-328, however, the lack of optical polarization (Tapia, personal communication in ref. 8) together with a rather cooler black-body emission suggests the UV emission is instead arising from an accretion disk. The absorption features further indicate that the whole binary system is immersed in a strong wind emanating from the system. These features make it far more similar to other X-ray cataclysmic variables (which also show no polarization⁸) than to AM Her, and as such emphasize the rather unusual nature of AM Her and 2A0311-227.

We thank the SRC for support, the IUE staff for observational assistance and Dr G. Bromage for helpful advice.

Received 13 November 1980; accepted 16 January 1981.

1. Tapia, S. *Astrophys. J. Lett.* **212**, L125 (1977).
2. Stockman, H. S. *et al. Astrophys. J.* **217**, 815 (1977).
3. Raymond, J. C. *et al. Astrophys. J. Lett.* **230**, L95 (1979).
4. Hearn, D. R. & Richardson, J. A. *Astrophys. J. Lett.* **213**, L115 (1977).
5. Coe, M. J. *et al. Nature* **279**, 509 (1979).
6. Swank, J. *et al. Astrophys. J. Lett.* **216**, L71 (1977).

7. Griffiths, R. E. *et al. Astrophys. J. Lett.* **232**, L27 (1979).
8. Charles, P. A. *et al. Astrophys. J. Lett.* **231**, L131 (1979).
9. Boggess, A. *et al. Nature* **275**, 372 (1978).
10. Warner, N. *Mon. Not. R. astr. Soc.* **191**, 43P (1980).
11. Heap, S. R. *et al. Nature* **275**, 385 (1978).
12. Hill, P. W. *IAU Circ. No.* 3492 (1980).
13. Bath, G. T., Pringle, J. E. & Whelan, J. A. J. *Mon. Not. R. astr. Soc.* **190**, 185 (1980).
14. Fabbiano, G. *et al. Preprint No.* 80-306 (Center for Astrophysics, 1980).
15. Tapia, S. *IAU Circ. No.* 3327 (1979).

Why is the central gas disk of the Galaxy tilted?

Leo Blitz*, James W.-K. Mark*†
& Rameshwar P. Sinha‡

* Astronomy Laboratory, University of California, Berkeley, California 94720, USA

† Lawrence Livermore Laboratory, Livermore, California 94550, USA

‡ National Radio Astronomy Observatory, Charlottesville, Virginia 22901, USA

Independent analyses of the HI and CO within 2 kpc of the galactic centre have shown that the gas distribution is best described as a disk-like structure with a rotation axis inclined 20–25° to that of the Galaxy^{1–3}. Detailed models of the gas as an expanding circular disk² or a thin elliptical bar³ describe the observations equally well, but they have no dynamical basis. We propose that the bending wave theory of Bertin and Mark⁴ can provide the dynamical explanation for the tilt, and that the predictions of the theory are independent of the kinematic model of the gas distribution. We show here that the predicted relationship between the disk surface density distribution and the height of the bending are satisfied for both the disk-like and bar-like models of the gas.

The original motivation for the theory^{4–7} was to provide an explanation for the warps in the outer parts of spiral galaxies, and it has satisfactorily explained several observations⁸. The model shows that the warp arises as a self-excited bending wave which is driven unstable because of its motion through the slowly rotating halo. Thus, the disk is rather like a 'flag waving in the wind (of the halo)'. The model described here to explain the tilt in the gas disk near the centre of our Galaxy is similar to this. In

this case, the gas disk moves through the more slowly rotating bulge and an analogous instability results in a 'flapping' of the disk. Direct observations of the bulge component of the galactic nucleus using OH masers related to Mira variables⁹, and IR radiation from population II stars¹⁰, indicate the presence of a slowly rotating spheroidal distribution of stars with a large scale height, comparable to that seen in other galaxies. Liszt and Burton model their CO and HI observations of the central regions of the Galaxy by a massive disk or bar ($> 10^9 M_{\odot}$) with a gaussian scale height of 100 pc and typical rotational velocities of 200–300 km s⁻¹. Thus, the conditions of a thin, rapidly rotating disk embedded within a larger, spheroidal mass distribution treated by the theory are applicable within the central 2 kpc of the Galaxy.

The bending wave theory⁵ predicts a scaling law

$$r^{1/2}\Sigma h \sim \text{constant} \quad (1)$$

where $\Sigma(r)$ and $h(r)$ are the surface mass density and height of bending (or tilt) as a function of radial distance. This scaling law arises because the bending wave propagates with nearly constant rate of outflow of wave energy through circles of radius r ,

$$2\pi r e c_g \sim \text{constant} \quad (2)$$

Equation (1) follows if we note that the wave energy density $\epsilon \propto \Sigma h^2$ and the group velocity $c_g \propto \Sigma$ (for detailed derivation see ref. 5). Although interactions with the bulge slowly amplify this rate of outflow, the amount of amplification per cycle of wave propagation is small and it is the short cycle time which determines the wave growth within the age of the Galaxy.

To examine the constancy of the product $\Sigma r^{1/2}h$, we rely on the detailed models of the nuclear disk for the geometric and kinematic properties of the disk. Burton and Liszt assume a constant surface density in their models and although there is excellent morphological agreement between the predicted and observed moment maps, the agreement between the contour gradients of their model and the observations is not particularly good. We assume that the observed contour gradients provide

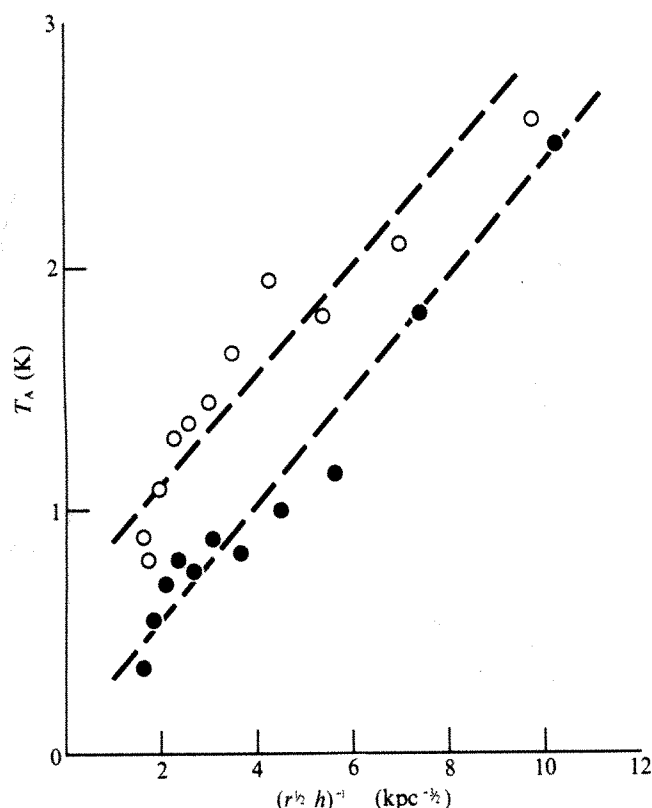


Fig. 1 Plot of HI antenna temperature (a quantity proportional to gas density) as a function of the quantity $(r^{1/2}h)^{-1}$ for the circular model (●) and bar-like model (○) described in the text. The dashed lines are least squares fits to the data.

information on the actual run of density of the disk. The observed antenna temperature (T_A) along the line $b = -l \tan 22^\circ$ should provide a good indication of the run of surface density because: (1) This locus of points is the axis of kinematic symmetry (corresponding very nearly to the projected major axis of the circular disk model) and observations along it provide the greatest contrast between the gas in the nuclear disk and the Galaxy at large. (2) The gas is surely optically thin and thus the antenna temperature of the gas should be directly proportional to the volume density. (3) The Burton and Liszt models show good agreement with the observations on the assumption of a disk of constant thickness. If this assumption is approximately correct, the observed HI antenna temperature should give the run of surface density.

To determine the surface density, we used the improved HI observations³ along the line $b = -l \tan 22^\circ$. This line intersects the midplane of the disk-like and bar-like models along a unique locus of points, and we determined the apparent LSR velocity along the locus for both models at 0.5° intervals. From the moment map of the HI observations, we found the HI antenna temperature on both sides of $b = 0$ and averaged the results. Because the data are satisfactorily modelled by a planar distribution with a constant inclination angle the product $\Sigma r^{1/2}h \propto T_A r^{3/2}$ for both the disk-like and bar-like models.

In Fig. 1 we plot T_A as a function of $(r^{1/2}h)^{-1}$ for both models. The points do not differ markedly from a straight line in either case, as expected from the theory. The dashed lines are the least squares fit to the data for each model. The data plotted in Fig. 1 include the HI emission for $l \geq 2^\circ$. Observable CO emission extends to nearly $l = 2^\circ$, and at lower longitudes the H_2 mass is likely to dominate the surface density of the nuclear disk. It is not possible to determine the H_2 mass from CO observations in the nuclear disk with any confidence, and we have therefore excluded the region of $l < 2^\circ$ from our analysis.

It seems reasonable to interpret Fig. 1 as implying that the run of density in the nuclear disk agrees with the predictions of the bending wave theory independently of the kinematics of the model used to describe the data. The largest uncertainty probably comes from the assumption of a constant thickness of the gas layer. The disk-like model could have an increasing scale height due to decreased self-gravity of the disk whereas the bar-like model may have the opposite tendency if it is really a triaxial ellipsoid rather than a thin disk. These possibilities can be tested only with better observations and more refined model making. Nevertheless, for the models we consider, which satisfactorily describe the present observations, the bending-wave theory provides a plausible dynamical basis for the tilt of the nuclear disk.

Received 3 November; accepted 24 December 1980.

1. Sinha, R. P. thesis, Univ. Maryland (1979).
2. Burton, W. B. & Liszt, H. S. *Astrophys. J.* **225**, 815–842 (1978); **226**, 790–816 (1978).
3. Liszt, H. S. & Burton, W. B. *Astrophys. J.* **236**, 779–797 (1980).
4. Bertin, G. & Mark, J. W.-K. *Astr. Astrophys.* **88**, 289–297 (1980).
5. Hunter, C. & Toomre, A. *Astrophys. J.* **155**, 747–776 (1969).
6. Kulsrud, R. M., Mark, J. W.-K. & Caruso, A. *Astrophys. Space Sci.* **14**, 52–54 (1971).
7. Mark, J. W.-K. *Astrophys. J.* **169**, 455–475 (1971).
8. Lake, G. & Mark, J. W.-K. *Nature* **287**, 705–706 (1980).
9. Baud, B. thesis, Univ. Leiden (1978).
10. Maihara, T., Oda, N., Sugiyama, T. & Okuda, H. *Publ. astr. Soc. Jap.* **30**, 1–19 (1978).

Radio acoustic measurement of fog-capping thermal inversions

G. Bonino*†, P. P. Lombardini†, A. Longhetto*‡
& P. Trivero*†

* Istituto di Fisica Generale dell'Università, Torino, Italy

† Istituto di Cosmo-geofisica del CNR, Torino, Italy

‡ Centro di Ricerca Termica e Nucleare dell'ENEL, Milano, Italy

The radio acoustic sounding system (RASS) is a remote technique for measuring temperature profiles in the lower troposphere^{1,2}. In essence, it consists of a powerful acoustic source beaming a short burst of sinusoidal waves towards the zenith. The speed of this pulse in its upwards path is proportional at every height to the square root of the local temperature. This speed is continuously measured from the ground by means of a Doppler radar. The radar echo is due to the change in refractive index of air compressed by the acoustic wave. The faint echo is enhanced by choosing an acoustic wave in Bragg resonance with the radio wave. The record of the measured sound speed displayed as a function of delay from the start of the acoustic beam leads to the acquisition of the temperature vertical profile. We now describe the use of such a system at metre-wavelengths in the study of fog-capping thermal inversions at a site in the Po Valley.

Some of our measurements with such a system have been published elsewhere^{3,4}. Comparison with other tropospheric soundings shows that RASS can give temperature profiles between 170 and 1,000 m with an accuracy of $\sim 0.2^\circ\text{C}$ (independent of height) and with a vertical resolution ~ 20 –30 m. The system has the advantages expected of remote sensing devices—short time resolution, rapid measurement and small cost of operation. We have made a particular study of fog-capping temperature inversions because of their importance in determining the fate of airborne pollutants discharged in the thermal plumes from tall industrial chimney stacks.

Since the Summer of 1979, our system has operated on the premises of the thermo-electric power plant of Turbigo. During September 1979, the fourth field experiment on atmospheric

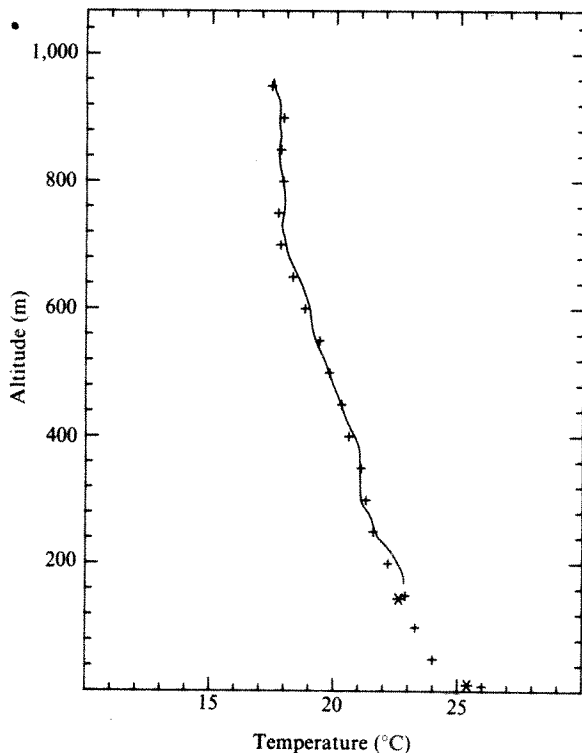


Fig. 1 Comparison of thermal profiles, 13 September 1979, 13 h 30 min: —, RASS; +, radiosonde on disposable balloon (Joint Research Centre); *, measurements at 2 and 145 m height along a smokestack of the power plant.

remote sensing, was carried out by 14 European teams under the auspices of the CEC (Commission of the European Communities) Environmental Research Program, in Turbigo⁵, using radiothermosondes, towed by disposable balloons, tethered balloons, model aircraft, to measure vertical air temperature profiles more or less simultaneously.

The thermal profiles obtained were compared with RASS soundings⁴ as shown in Fig. 1, which includes a temperature profile obtained by the radiothermosonde. Figure 1 shows that RASS data compare favourably with other readings; any local disagreement in the plots is within 0.5 °C, and may be justified by differences in the footprint of the sounding, a few kilometres, and/or the time of measurement. Approximately 100 soundings were compared during the CEC campaign. Limitations associated with this technique are essentially due to strong horizontal winds and large convective vertical motions^{1,6}. Wind shears associated with strong wind distort the acoustic burst which limit the applications of this technique worsens the coherent matching between radio and acoustic waves and causes a lowering of the sounding range. This is not a major drawback in that such conditions ensure a very effective stirring of air and the temperature profile is close to the 'adiabatic lapse rate', which corresponds to a uniform vertical temperature decrease of 0.01 °C m⁻¹.

The temperature error Δt (°C), arising from the effect of the vertical component w (m s⁻¹) of the wind on the RASS measurements, is given by $\Delta t = -1.7 w$. Although this error could cause trouble when evaluating air thermal structure associated with diurnal convective unstable layers, it can be eliminated by using an acoustic radar technique to measure the wind vertical component of velocity independently^{7,8}, or it can be reduced by averaging temperature measurements over suitable periods of time.

In the fog conditions, discussed here, however, these limitations are negligible.

The Turbigo site is located in the centre of the Po Valley, and is characterized by prevailing low wind speed regimes. A non-

uniform thermal structure of the planetary boundary layer (PBL) frequently associated with this meteorological situation, may lead, for example, to nocturnal stable inversion layers, subsidence inversions and fog layers due to water vapour condensation processes. For these cases no prognostic model exists⁹. Consequently, more reliable data from PBL observation are needed, for which knowledge of the vertical thermal profile is important.

One of the most difficult structures of the PBL to measure is the radiative fog layer capped by an elevated thermal inversion of unknown depth and intensity.

We need to know: (1) the thickness of fog layer; (2) the intensity of radiative thermal inversion at the top of fog; (3) the atmospheric stability parameter above the fog-capping inversion, at height 500–1,000 m. SODAR may reveal the first data. The second or third data have to be assessed with disposable radiosondes. However, economy and air traffic regulations impose severe limitations on the latter method.

The present experiment refers to synoptic weather conditions marked by fog in the Po Valley in connection with a high levelled pressure over the Mediterranean area, between 8 and 14 February 1980. On the last day the high pressure promontory weakened, and fog disappeared. Figure 2 and Table 1 show the RASS soundings, balloon soundings from Linate airport and temperature readings along a Turbigo smokestack. The available data from the Linate radiosoundings are only 1–2 between 0 and 1,000 m because these measurements are used for meteorological forecasts and for information regarding air traffic.

Figure 2 shows that: RASS soundings describe the internal thermal structure of PBL up to a high of ~1 km; RASS thermal profile is continuous both in time and in height; vertical resolution and temperature accuracy are satisfactory.

The difference between the thermal profiles on 8 February as deduced by the available Linate radiosonde data and by RASS soundings is interesting. This radiosonde profile would have brought the observer to assume a fairly stable atmosphere above

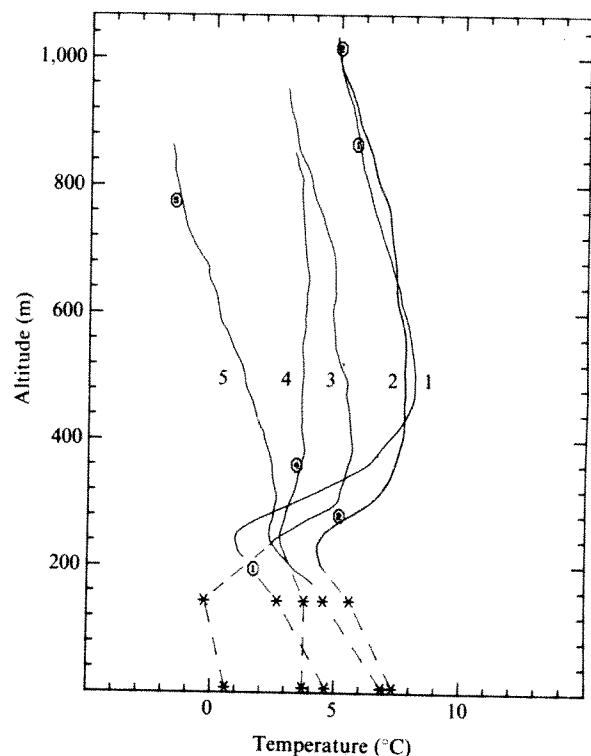


Fig. 2 Atmospheric temperature vertical profile: —, RASS measurements; O, Linate airport sounding data referred to n th sounding in Table 1; *, measurements at 2 and 145 m height along a smokestack of the power plant; ---, interpolation line connecting experimental points.

Table 1 Atmospheric profiles measured with RASS

RASS profile	Date (February 1980)	Time (p.m.)	Weather conditions	Visibility at Linate airport at 3.00 p.m.	R_H (%)
1	8	4.30	Foggy	400 m	98.7
2	9	4.20	Foggy	600 m	86.8
3	12	12.00	Foggy	1,000 m	97.8
4	13	12.00	Foggy	800 m	98.8
5	14	12.00	Not foggy	4,500 m	78.5

The Linate sounding is always made at 1.00 p.m.

the fog layer, from 300 to 1,000 m. RASS soundings show however, that the stability was not uniformly distributed: a very strong stability between 300 and 400 m being surmounted by a near neutral stratification.

Clearly, such a different representation of the vertical structure of air stability above the fog layer strongly influences pollution forecasts. The radiosonde profiles would have predicted that smoke released by large hot emissions from an elevated point source could be diluted in the higher regions of the PBL. The RASS profile, however, could predict trapping in the capped inversions and hence the danger of a high pollution level.

This example stresses the importance to environmental management of having a reliable vertical thermal profile in real time. RASS has qualities which could fulfil this role.

Received 4 July; accepted 31 December 1980.

1. North, E. M., Peterson, A. M. & Parry, H. D. *Bull. Am. met. Soc.* **54**, 912-919 (1973).
2. Frankel, M. S. & Peterson, A. M. *Radio Sci.* **11**, 157-166 (1976).
3. Bonino, G., Lombardini, P. P. & Trivero, P. *IEEE Trans. Geosci. Electron.* **GE-17**, 179-181 (1979).
4. Bonino, G., Lombardini, P. P. & Trivero, P. *Nuovo Cimento* **3C**, 207-214 (1980).
5. Guillot, P. *Proc. 10th NATO/CCMS int. Tech. Meet. Air Pollution Modeling*, Rome, 85-94 (1979).
6. North, E. M. *Final rep., SU-SEL-73-021*, 108 (Stanford Electronics Laboratories, Stanford, 1974).
7. Beran, D. W., Little, C. G. & Willmarth, B. C. *Nature* **230**, 160-162 (1971).
8. Spizzichino, A. *J. geophys. Res.* **79**, 5585-5591 (1974).
9. Yu, T. W. *J. appl. Met.* **17**, 28-33 (1978).

Low frequency 0.045-Hz swell from the Southern Ocean

F. A. Shillington

Department of Oceanography, University of Cape Town, Rondebosch 7700, South Africa

Unusual, low-frequency, one-dimensional energy spectra of surface gravity waves from measurements with a buoy in deep water (167 m) near Cape Town, South Africa, are discussed here. Classical treatment of the time series of spectra showed that the wave source was 2,800 nautical miles distant and probably in the Southern Ocean. Examination of South African Weather Bureau surface pressure analyses of the South Atlantic depict an intense, slow moving cyclone, with geostrophic wind velocities >50 knots at the same time and place as the origin of the waves. The wave spectra are highly unusual in that they reveal quasi-regular swell at very low frequencies with spectral energy densities an order of magnitude larger than those normally present at these frequencies. The wave event and its associated meteorological conditions are also discussed.

Surface gravity waves have been measured on a regular basis from positions near Cape Town, South Africa, for several years¹⁻³. These sites are open to waves from the South Atlantic and Southern Ocean. Because long fetches are possible to the south and west of South Africa, it is accepted that the peak

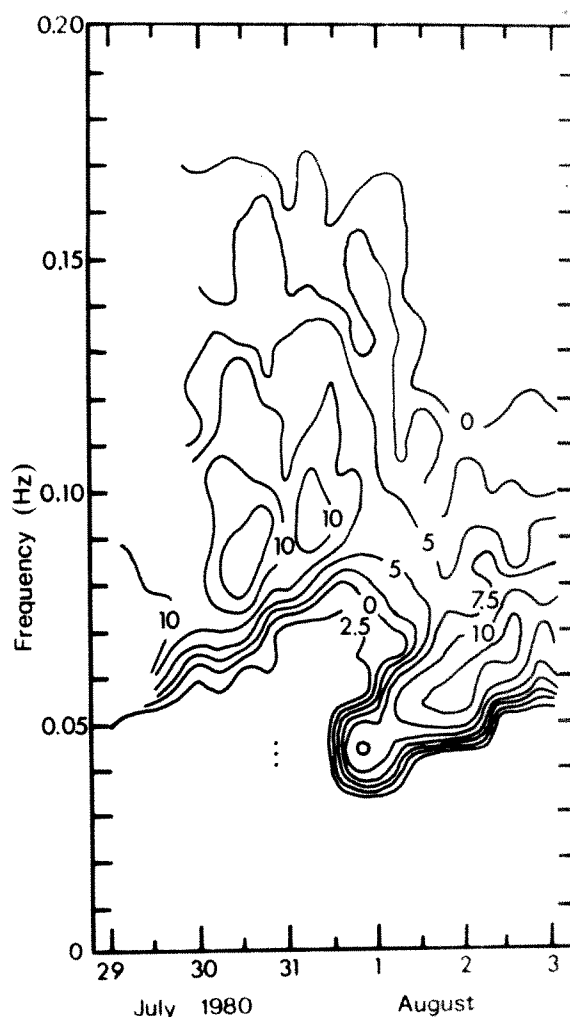


Fig. 1 Energy-frequency-time diagram with days marked at 12 h GMT and wave energy contours in dB above $1 \text{ m}^2 \text{ s}$.

frequency of the wave spectra is often low (between 0.06 and 0.10 Hz). However, it is rare to measure significant wave energy in frequency bands with the central frequency lower than 0.05 Hz (period 20 s). This is the third such occurrence (see refs 1, 2) in ~8 yr of measurements. The signature of the event in the energy-frequency-time space and the local height of these low frequency waves rules out the explanation that the storm has originated in the Northern Hemisphere as do the Saint Helena rollers described by Cartwright *et al.*⁴.

The wave recordings were made with a Datawell waverider buoy every 6 h for 20-min intervals. Two buoys are maintained: one in a depth of 167 m of water and the other in 20 m of water. The contours of energy in frequency-time space were similar for both buoys. Only the measurements from the deep buoy are presented, as these are free from bottom effects. Spectral values were computed by standard FFT techniques from 17-min long time series sampled every 0.5 s. The estimates have 20 d.f. with 80% confidence levels of 0.62 and 1.42 times the estimates. Frequency resolution is 0.0098 Hz. The spectra have been plotted to give an energy-frequency-time plot (see Fig. 1), with contours of wave energy density at logarithmic intervals.

Assuming that the storm-wave generation area is small compared with its distance from the recorder, classic treatment of the spectra shows that there will be a linear increase of swell frequency with time. If t_0 is the time of origin of such swell, the distance, D , of the generating area can be determined by extrapolating the slope ($\Delta f / \Delta t$) of the 'ridge' in Fig. 1 back to the zero frequency axis and noting that the group velocity of waves

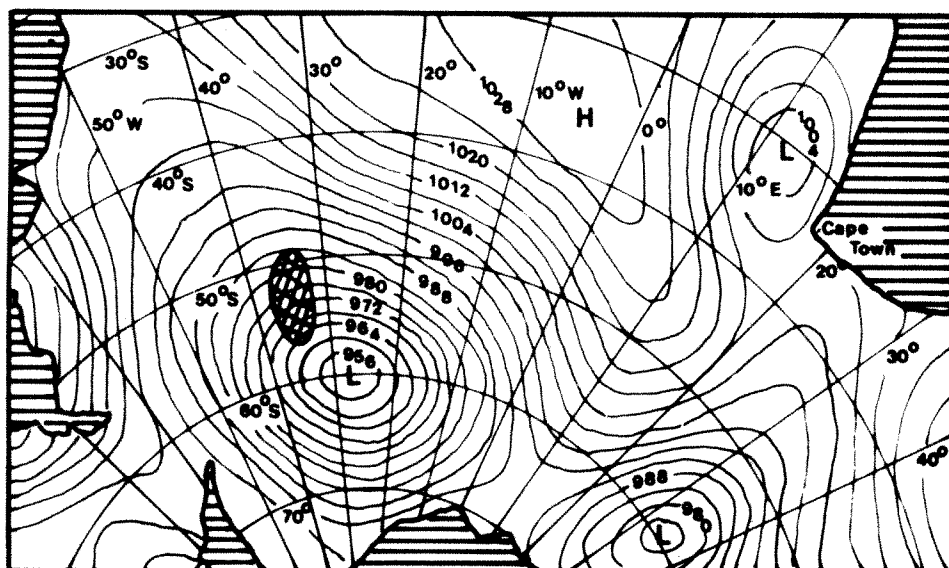


Fig. 2 The South African Weather Bureau synoptic surface pressure analysis for 12 h GMT 29 July 1980. Contours are at 4-mbar intervals and the probable fetch for the waves is hatched. The map projection is polar stereographic.

of frequency f is

$$Cg = \frac{g}{4\pi f}$$

so that

$$D = g / \left(4\pi \cdot \frac{\Delta f}{\Delta t} \right).$$

Using this technique, we find that the wave source origin was at 00 h GMT, 29 July 1980, 2,800 nautical miles from the recorder (Cape Town). The lowest frequency identifiable as a definite peak is at 0.045 Hz; these waves travel at 800 nautical miles per day and reached the wave recorder at 12 h GMT on 1 August 1980. Slower moving 0.06 Hz waves arrived one day later at 12 h GMT.

The normal source area of wave events directed towards Cape Town is the South Atlantic. Examination of the South African Weather Bureau synoptic forecast weather surface pressure charts revealed an intense depression centred at 55°S, 35°W at 00 h GMT, 29 July 1980 (see Fig. 2). The central pressure of this cyclone deepened from 964 to 956 mbar in the 24 h from 12 h 28 July to 12 h 29 July. A further deepening of the low to 948 mbar occurred on 30 July. The storm track followed a course to the south-east with the centre being at 68°S on 1 August 1980 (see Fig. 3). At this time, winds within 1,000 miles of Cape Town were from the south-east and unfavourable for the generation of long period waves. Geostrophic wind-speeds higher than 50 knots are evident on the chart (Fig. 2).

What windspeed is required to generate waves with a peak frequency of 0.045 Hz? The present method follows Pierson and Moskowitz⁶ who showed that for a fully developed sea, the peak frequency (f_m) and windspeed are (U) related by

$$f_m = 1.37 U^{-1} \quad (U \text{ measured at 10 m height, in } m s^{-1}, f_m \text{ in Hz}).$$

For $f_m = 0.045$ Hz, $U = 30.4 m s^{-1}$ that is, a 60-knot wind is required to generate a fully developed spectrum with a peak frequency of 0.045 Hz.

As a fully developed spectrum was not measured, the present calculation is based on the work of Hasselmann *et al.*⁷ to find the size of fetch that would produce a wave spectrum with a peak at $f_m = 0.045$ Hz. The empirical equation found from the JONSWAP⁷ data relates the dimensionless peak frequency $\tilde{f}$ to the dimensionless fetch $\tilde{x}$ by

$$\tilde{f} = 3.5 \tilde{x}^{-0.33}$$

where

$$\tilde{f} = \frac{f_m U}{g} \quad \text{and} \quad \tilde{x} = \frac{xg}{U^2}$$

For $U = 30 m s^{-1}$, $\tilde{f} = 0.138$ and $\tilde{x} = 1.64 \times 10^4$. The length of the fetch is 1,500 km. These values place the results at the limit of the JONSWAP data set⁷.

Using $\tilde{x}$ above and the work of Mitsuyasu and Rikiishi⁸, the minimum duration required to generate a spectrum with a peak frequency of 0.045 Hz with a fetch of 1,500 km, is 23 h. Again,

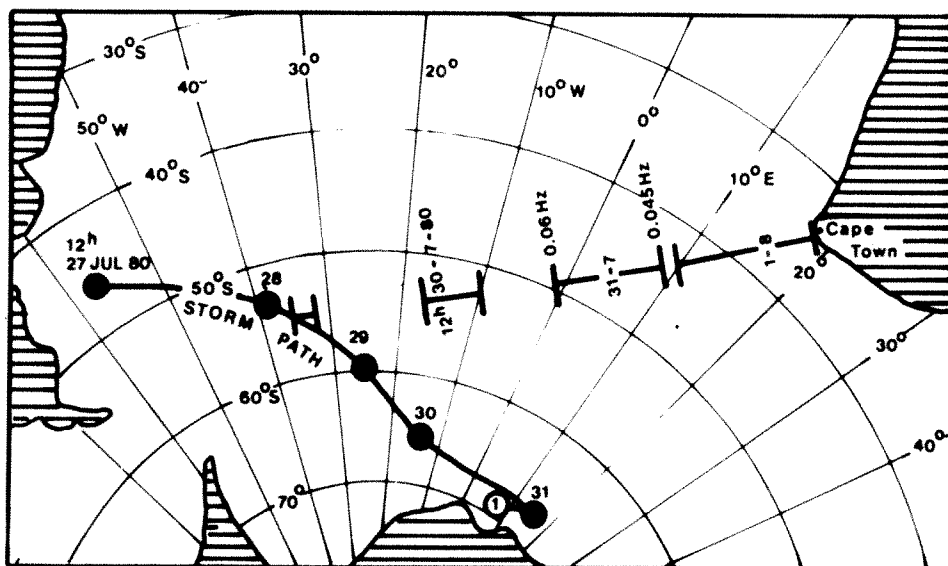


Fig. 3 Chart with storm track and storm centre locations together with wave travel time and positions of 0.045 Hz and 0.06 Hz components.

because the work cited⁸ is for small dimensionless fetches, extrapolation to larger fetches might be dangerous.

If these large orders of magnitude for the windspeed, fetch and duration are required to generate the observed wave conditions ($U = 30 \text{ m s}^{-1}$, fetch = 1,500 km, duration = 23 h), then even if it was not unusual to find deep, low pressure cyclones in this part of the Southern Ocean⁹, it would be exceptionally rare to get a fetch of the order of 1,000 km with sustained 60-knot winds for a period of a day, generating waves in one particular direction. Careful examination of one year's wave spectra recorded at South Uist in the Outer Hebrides¹⁰ at latitude 57°N, longitude 8°W, shows that the spectral energies reported here were a factor of two greater at 0.045 Hz than any of those at South Uist. The dangers of extrapolating measurements^{7,8} beyond their respective domains are accepted, but with the limited data available, it seemed to be the only logical approach.

I thank the Council for Scientific and Industrial Research's National Research Institute for Oceanology, Stellenbosch and the Fisheries Development Corporation for maintenance of equipment and availability of the deep sea data. The work was funded in part by the South African Committee for Oceanographic Research.

Received 16 October; accepted 13 January 1981.

1. Shillington, F. A. & Harris, T. F. W. *Deut. hydrogr. Z.* **31**, 67–81 (1978).
2. Harris, T. F. W., Marshall, J. N. & Shillington, F. A. *Nature* **224**, 154 (1973).
3. Shillington, F. A. & Britten-Jones, A. S. *Afr. J. Sci.* **75**, 495–500 (1979).
4. Cartwright, D. E., Driver, J. S. & Tranter, J. E. *Q. J. R. met. Soc.* **103**, 655–683 (1977).
5. Munk, W. H. *et al. Phil. Trans. R. Soc. A* **255**, 505–584 (1963).
6. Pierson, W. J. & Moskowitz, L. J. *geophys. Res.* **69**, 5181–5190 (1964).
7. Hasselman, K. *et al. Ergebn. Deut. hydrogr. Z.* **A12**, (1973).
8. Mitsuyasu, H. & Rikishi, K. *J. Fluid Mech.* **85**, 705–730 (1978).
9. Newton, C. W. (ed.) *Met. Monogr.* **13** (1972).
10. Fortnum, B. C. H. *et al. Inst. Oceanogr. Sci. Rep. No.* 71 (1979).

Changes of ^2H and ^{18}O enrichment of meteoric water and Pleistocene glaciation

Russell S. Harmon

Isotope Geology Unit, Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

Henry P. Schwarcz

Department of Geology, McMaster University, Hamilton, Ontario, Canada L8S 4M1

Recent theoretical modelling¹ has suggested that the well documented δD – $\delta^{18}\text{O}$ relationship for modern meteoric waters may have changed during the Pleistocene glacial periods in response to increased ocean surface air humidity. Here we present isotopic data on fluid inclusions and host speleothems which support this hypothesis. Depositional temperatures can be determined for speleothems formed in equilibrium conditions from the temperature dependence of the calcite–water ^{18}O fractionation by isotopic analysis of both entrapped fluid inclusions and host calcite. Palaeotemperatures for five areas of east-central North America and Bermuda, calculated assuming the present δD – $\delta^{18}\text{O}$ meteoric water relationship for fluid inclusion waters, are observed to be too low during late Pleistocene glacial periods, in some instances falling below 0 °C. By contrast, interglacial palaeotemperatures are largely equivalent to those in the areas at present. As speleothem deposition cannot occur at subzero temperatures, a possible solution to this dilemma is a shift in the intercept of the meteoric water relationship.

The stable isotopic composition of most coastal and continental meteoric waters not subjected to extensive evaporation form a linear array (Fig. 1a) best described by the relationship:

$$\delta\text{D} = X\delta^{18}\text{O} + d_0 \quad (1)$$

where $X = 8.0$ and d_0 , the so-called deuterium deficiency, is +10‰ (refs 2, 3). This correlated behaviour of D and ^{18}O variations in the hydrosphere results from both equilibrium and kinetic effects which occur during the processes of evaporation and condensation in the atmosphere^{1–4}. Liquid–vapour fractionations are such that vapour evaporated from the ocean is proportionally depleted in D and ^{18}O . Water subsequently condensed in the atmosphere will be enriched in both heavy isotopes to an extent determined by the isotopic composition of the vapour source and the temperature at the time of condensation. Thus, precipitation from an air mass moving away from its tropical oceanic source towards higher latitudes will become progressively depleted in D and ^{18}O , producing a pronounced and characteristic geographic distribution in the isotopic composition of continental precipitation (Fig. 1b). This natural evaporation–condensation cycle can be modelled in terms of a Rayleigh distillation process, the slope of the δD – $\delta^{18}\text{O}$ relationship in equation (1) being characteristic of liquid–vapour equilibrium^{1–3} and the intercept d_0 , indicative of the extent to which kinetic processes have affected the system.

At present this general, worldwide δD – $\delta^{18}\text{O}$ relationship is determined by the isotopic composition of the ocean, the ultimate source of all continental precipitation, as well as the prevailing climatic conditions in which evaporation and condensation occur. During the last glacial maximum at 18,000 yr BP the ocean was enriched in ^{18}O by ~1.1–1.4‰ (ref. 5); the corresponding shift in δD is undetermined at present, but was presumably close to +10‰. If prevailing climate conditions change so as to amplify or diminish kinetic fractionation during evaporation or condensation, d_0 will be altered. The data we present here suggest that this may have been the case during the Pleistocene ice ages.

Oxygen isotope variations in calcite speleothems can yield a relative record of past climatic fluctuations if deposition occurred in conditions of isotopic equilibrium with their parent seepage waters^{6–8}. Moreover, if the ^{18}O content of the parent seepage can be estimated from analysis of fluid inclusion trapped within a speleothem at the time of deposition⁹, then the temperature of deposition can be calculated from the temperature dependence of the calcite–water ^{18}O fractionation as given by O'Neil *et al.*^{10,11}:

$$10^3 \ln \alpha_{c-w} = 2.78 \times 10^6 T^{-2} + 2.89 \quad (2)$$

where $\alpha_{c-w} = (\delta^{18}\text{O}_{\text{calcite}} + 1,000) / (\delta^{18}\text{O}_{\text{water}} + 1,000)$. Inasmuch as the temperature deep within a cave is related to the long-term, average temperature on the surface above the cave, such isotopic palaeotemperatures should reflect surface temperature variations throughout the period over which the speleothem was deposited.

Direct analysis of the ^{18}O composition of fluid inclusion waters has not yet been satisfactorily demonstrated, the two major problems being the possibility of continued isotopic exchange between the host calcite and water subsequent to entrapment and that of exchange induced during crushing and extraction for analysis. It has previously been shown, however, that the D and ^{18}O contents of modern cave seepage in temperate areas of North America and western Europe are equivalent to the surface precipitation from which the ground-water seepage was derived^{9,12,13} (Fig. 1a). It should therefore be possible using equation (1) to estimate the $^{18}\text{O}/^{16}\text{O}$ ratio of speleothem fluid inclusion waters from an analysis of the D/H ratio. Table 1 shows the observed agreement both between the $^{18}\text{O}/^{16}\text{O}$ ratios calculated for fluid inclusion waters from modern

Table 1 Isotopic data and calculated isotopic temperatures for modern soda-straw-seepage waters and stalagmite-fluid inclusion pairs for sites in east central North America and Bermuda

Locality	A Soda-straw-seepage water pairs		B Stalagmite-fluid inclusion pairs			Isotopic temperature (°C)		Measured cave temperature (°C)
	$\delta^{18}\text{O}_c$	$\delta^{18}\text{O}_w$	$\delta^{18}\text{O}_c$ (‰SMOW)	δD_n	$\delta^{18}\text{O}_n$	T_A	T_B	
Iowa	24.10 (6)	-7.77 (9)	24.10 (1)	-55.6 (2)	-8.19	10.6	9.0	8.8
West Virginia*	23.15 (7)	-8.68 (8)	23.64 (4)	-60.9 (4)	-8.86	10.7	8.0	10.8
Kentucky	25.81 (8)	-5.59 (8)	—	—	—	12.9	—	13.5
Texas	26.78 (7)	-3.44 (11)	—	—	—	18.1	—	20
Bermuda	27.59 (6)	-2.07 (9)	—	—	—	20.8	—	21.0

$\delta^{18}\text{O}_c$ is measured speleothem calcite $^{18}\text{O}/^{16}\text{O}$ ratio.

$\delta^{18}\text{O}_n$ is calculated fluid inclusion $^{18}\text{O}/^{16}\text{O}$ ratio from equation (1).

$\delta^{18}\text{O}_w$ is measured water $^{18}\text{O}/^{16}\text{O}$ ratio.

δD_n is measured fluid inclusion D/H ratio.

Isotopic temperature is temperature calculated from equation (2).

Number of samples analysed are given in parenthesis.

* Data from ref. 28.

stalagmites and that measured for cave seepage waters and also between calculated isotopic temperatures for these speleothem-fluid inclusion pairs and measured cave temperatures. In both cases the differences are just slightly larger than the limits of the analytical uncertainty. Using the procedures described in ref. 14, D/H ratio measurements for speleothem fluid inclusions

have a precision of about $\pm 3\%$, whereas speleothem calcite $^{18}\text{O}/^{16}\text{O}$ ratio determinations have a precision of $\pm 0.1\%$. This translates to an overall uncertainty of about ± 2 to 3°C in estimated palaeotemperatures, given the inherent small-scale inhomogeneities in the volume of speleothem material used for an individual analysis. This uncertainty, though substantial, is

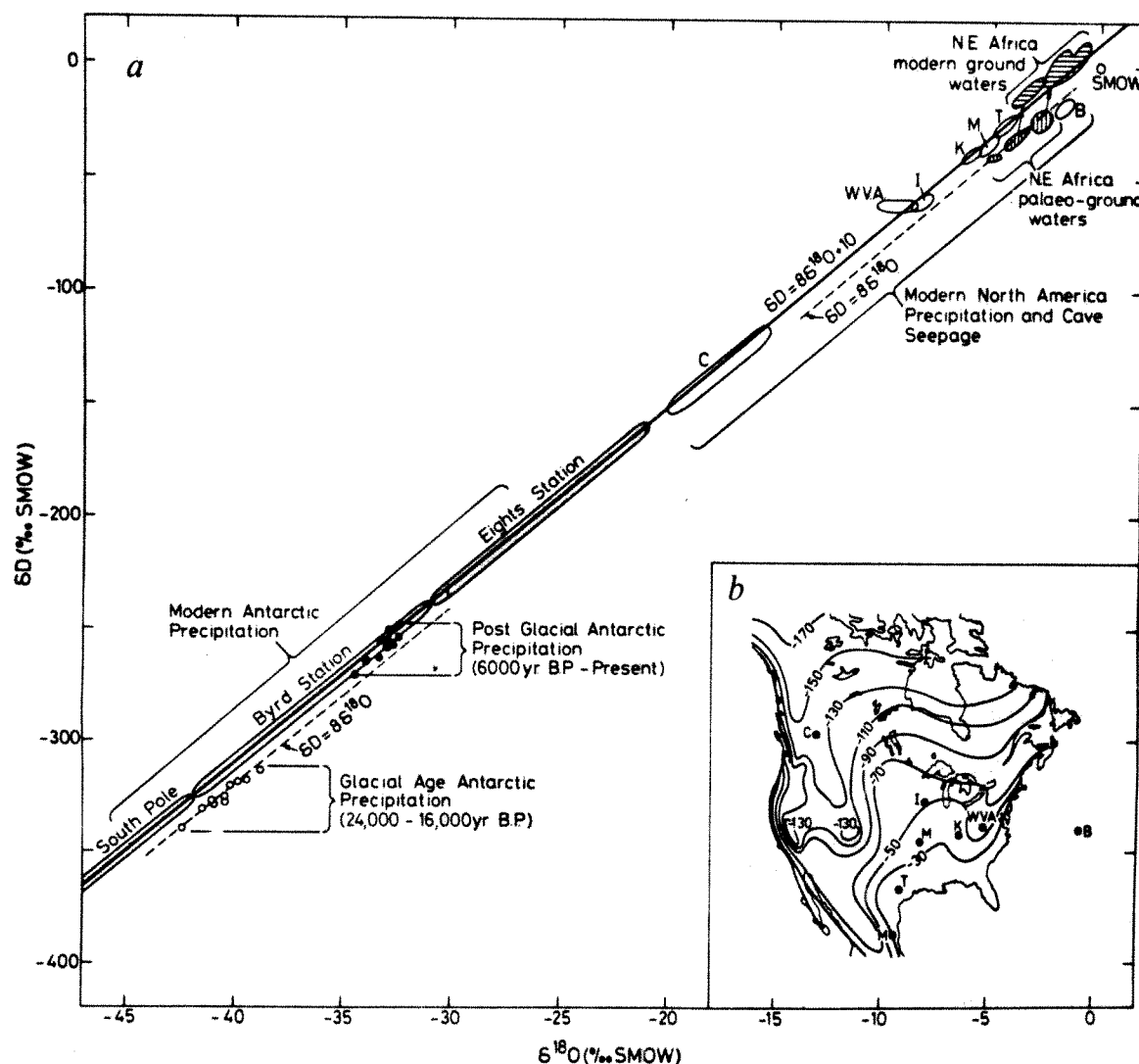


Fig. 1 a, Relationships between D/H and $^{18}\text{O}/^{16}\text{O}$ variations in various modern meteoric waters^{2,3}. Specific fields shown are those for modern Antarctic precipitation^{20,21}, post-glacial and glacial age ice from Byrd Station, Antarctica¹⁹, modern seepage waters in caves of east-central North America^{9,12,15} modern and palaeo groundwaters in North Africa²²⁻²⁴ and modern seawater^{25,26}. Modern meteoric waters are best described by the relationship $\delta\text{D} = 8\delta^{18}\text{O} + 10$ whereas that for glacial-age meteoric waters appears to be $\delta\text{D} = 8\delta^{18}\text{O}$. b, Map of North America showing approximate δD contours for meteoric surface waters (after Taylor)²⁷. Also shown by the North American sample localities discussed in the text.

Table 2 Isotopic composition, ages and calculated depositional temperatures for speleothem–fluid inclusion pairs from North American sites

Sample no.	Age (kyr)	$\delta^{18}\text{O}_c$	δD_w	$\delta^{18}\text{O}_w^*$ ‰ SMOW	$\delta^{18}\text{O}_w^{**}$	T^* (°C)	T^{**} (°C)	Climate type	ΔT (°C) ‡
Iowa									
74014:5	8.5	24.5	-67.1	-9.6	-8.4	2	7	t	-2
10		24.2	-72.8	-10.4	-9.1				
	11.5					0	5	t	-4
10		24.3	-72.0	-10.3	-9.0				
9	13.0	24.8	-72.0	-10.3	-9.0	-1	3	t	-6
8	14.0	24.7	-74.9	-10.6	-9.4	-3	2	t	-7
7	17.0	24.2	-79.4	-11.2	-9.9	-3	2	g	-7
4	20.0	24.1	-76.9	-10.9	-9.6	-1	3	g	-5
12	23.0	24.1	-72.3	-10.3	-9.0	1	5	g	-3
74025:1	2.5	24.1	-55.0	-8.1	-6.9	9	14	i	0
West Virginia									
GV2: 1,2	60	25.3	-62.3	-9.0	-7.8	1	6	t	-5
3	70	26.3	-61.9	-9.0	-7.7	-2	2	g	-8
4	97	23.8	-52.2	-7.8	-6.2	12	17	i	+1
5	104	23.8	-56.6	-8.3	-7.1	10	15	t(i)	+4
6,7,8	159	25.1	-59.2	-8.7	-7.4	3	8	g	-3
NB10:1	169	23.7	-65.0	-9.4	-8.1	6	11	g	0
2	172	23.5	-53.8	-7.9	-6.7	12	18	t	+6
3	174	23.0	-61.5	-8.9	-7.7	10	15	g	+4
4	197	23.6	-68.5	-9.8	-8.6	4	9	t	-1
5	200	23.8	-65.8	-9.5	-8.2	5	10	t	-1
Kentucky									
72041:10	106	24.9	-44.2	-6.8	-5.5	12	17	i	-2
12	112	25.7	-45.9	-7.0	-5.7	8	13	t	-1
11	122	24.8	-42.1	-6.5	-5.3	13	19	i	0
7	127	25.5	-61.6	-9.0	-7.7	1	6	g	-8
5	156	25.7	-49.2	-7.4	-6.2	6	11	g	-3
4	171	24.7	-45.0	-6.9	-5.6	12	18	t	+4
3	195	24.7	-48.2	-7.3	-6.0	11	16	t	+2
2	217	24.6	-39.1	-6.1	-4.9	16	21	i	+2
Texas									
72008:2	32	25.9	-51.8	-7.7	-6.5	4	9	g	-11
3	37	25.6	-64.5	-9.3	-8.1	-1	4	g	-16
4	41	26.3	-57.0	-8.4	-7.1	0	5	g	-15
5	45	26.0	-50.8	-8.4	-7.1	1	6	g	-14
7	60	26.7	-26.5	-4.6	-3.3	14	19	t	-1
Mexico									
71042:1	53	27.7	-14.5	-3.1	-1.8	16	22	t	-2
3	55	27.1	-37.8	-6.0	-4.7	6	12	t	-12
5	78	27.2	-32.7	-5.3	-4.1	9	14	t	-10
10	95	27.2	-37.5	-5.9	-4.7	6	11	t	-15
Bermuda*									
75001:4	25	28.5	-16.4	-3.4	-1.7	11	14	g	-1
75003:1	29	30.5	-19.2	-3.9	-2.2	2	5	g	-13
75004:12	99	29.2	-8.5	-2.1	-0.4	14	17	i	-7
11	101	28.2	-8.4	-2.1	-0.4	18	21	i	-3
13	104	28.5	-24.0	-4.7	-3.0	6	9	t	-8
9	109	27.5	-14.7	-3.2	-1.5	16	19	i	-5
7	112	29.0	-19.2	-3.9	-2.2	7	10	t	-7
4	115	28.2	-10.3	-2.4	-0.7	17	20	i	-4
5	118	28.0	-19.9	-4.0	-2.3	11	19	i	-10
73036:18	166	28.1	-20.1	-4.0	-2.4	6	9	g	-8
10	172	27.1	-11.0	-2.5	-0.8	21	24	t	+8
12	176	26.4	-17.0	-3.5	-1.8	19	22	g	+6
17	177	27.5	-24.0	-4.7	-3.0	10	13	g	-4
15	185	26.9	-24.1	-4.7	-3.0	12	15	g	-2

Age is interpolated age estimated from $^{230}\text{Th}/^{234}\text{U}$ dates on two or more horizons within the speleothem, assuming constant linear growth rate. $\delta^{18}\text{O}_c$ is measured oxygen isotopic composition of calcite. δD_w is measured hydrogen isotopic composition of fluid inclusion water. $\delta^{18}\text{O}_w^*$ is calculated oxygen isotope ratio of inclusion water assuming $\delta\text{D}_w = 8\delta^{18}\text{O}_w + 10$. $\delta^{18}\text{O}_w^{**}$ is calculated oxygen isotope ratio of inclusion water assuming $\delta\text{D} = 8\delta^{18}\text{O}_w$. T^* is isotopic temperature calculated from $\delta^{18}\text{O}_c$ and $\delta^{18}\text{O}_w^*$ using equation (2). T^{**} is isotopic temperature calculated from $\delta^{18}\text{O}_c$ and $\delta^{18}\text{O}_w^{**}$ using equation (2).

† For Bermuda we have assumed that for modern and interglacial precipitation: $\delta\text{D} = 6\delta^{18}\text{O} + 4.2$ while for transitional and glacial times: $\delta\text{D} = 6\delta^{18}\text{O}$ (see refs 9, 13).

‡ A climatic classification of the samples was constructed from a correlation of their ages with the foraminiferal isotopic record of deep-sea core V23-238⁵ for which a time scale has been established from the position of the ~730,000 yr BP. Brunhes–Matuyama magnetic reversal and the assumption of a constant sedimentation rate throughout the core. $\delta^{18}\text{O}$ values of foraminifera ($\delta^{18}\text{O}_f$) increase with increasing ice volume because of the increase in $\delta^{18}\text{O}$ of the ocean and lower oceanic temperatures during glacial periods. Thus speleothems deposited at times when $\delta^{18}\text{O}_f > -1.25\text{‰}$ are called glacial (g); when $\delta^{18}\text{O}_f < -1.70\text{‰}$ we have interglacial (i) conditions; between these limits we refer to the climate as intermediate (t).

smaller than the 4–14 °C temperature difference postulated from climate reconstruction for east-central North America during the 18,000 yr BP Wisconsinan glacial maximum¹⁵. This technique should, therefore, be applicable to ancient speleothems where it would have the potential to provide a proxy record of late Pleistocene terrestrial climate change which is otherwise difficult to obtain.

Isotopic data for 11 late Pleistocene speleothems deposited in equilibrium conditions within caves in east-central North America and Bermuda are given in Table 2. Ages for individual calcite–fluid inclusion pairs were interpolated from growth

curves established by $^{230}\text{Th}/^{234}\text{U}$ dating of multiple horizons within the speleothem¹⁶.

We have previously shown that fluid inclusions from five of these speleothems became depleted in deuterium as climate shifted from interglacial to glacial conditions¹⁵; the average shift in δD in north-central North America was about -12‰. Palaeotemperatures calculated using equation (1) for these five samples and six additional ones are presented in Table 2. While most palaeotemperatures range between modern values at a site and 0 °C, in some instances they fall below 0 °C (for example, Iowa) and in others are clearly too low considering the

geographic location of the cave sites (for example, Texas and Bermuda).

Groundwater recharge and calcite precipitation cannot occur when cave temperature is below freezing. Groundwater recharge is possible in certain permafrost conditions where the annual average surface temperature is below 0 °C, such as in the high Arctic at present¹⁷, and cave temperatures may be maintained above freezing in an area where the average surface temperature is subzero¹⁸. These situations, however, represent arctic and alpine anomalies unrelated to the temperate lowland and marine karst areas, under discussion here. Therefore, we consider the very presence of a speleothem deposit as evidence that the cave temperature was above the freezing point during the time of its deposition.

A possible resolution of this dilemma is that equation (1) changed during Pleistocene glacial periods. The general δD - $\delta^{18}O$ meteoric water relationship is geographically somewhat variable at present³. For example, precipitation in Japan and the Mediterranean is displaced to lower $\delta^{18}O$ values by ~1.5% relative to other meteoric waters with an equivalent δD value, while for most oceanic islands the relationship is $\delta D = 4.2\delta^{18}O + 0.1$. More significantly, temporal variation in the δD - $\delta^{18}O$ relationship was noted by Epstein *et al.*¹⁹ in their study of the ice core from Byrd Station, Antarctica. Here post-glacial precipitation falls on the $\delta D = 8\delta^{18}O + 10$ line as does modern precipitation^{20,21}, whereas glacial-age precipitation is best described by the relationship $D = 7.9\delta^{18}O$ (Fig. 1a). Our data are compatible with the idea that Pleistocene meteoric waters elsewhere also exhibited a relative deuterium deficiency during glacial periods. Thus during glacial periods, $X \approx 8$ and $d_0 \approx 0$ with d_0 approaching its present value of +10 during transitions from glacial to interglacial climate.

Independent evidence for such changes in d_0 is provided by hydrologic studies in North Africa²²⁻²⁴ where low- ^{14}C (Pleistocene-age) groundwaters have been observed with $d_0 = 0$, whereas for modern precipitation $d_0 \approx 10$ (Fig. 1a). It therefore seems that during at least the last glacial period of the Pleistocene $X \leq 8$ and d_0 was significantly less than 10; both variables acquiring the present value (8, 10) shortly after the present interglacial period began some 10,000 yr ago.

We have recalculated temperatures for speleothem samples deposited during glacial and transitional periods assuming the extreme boundary conditions of $X = 8$ and $d_0 = 0$ (Table 2). In so doing, all isotopic temperatures become positive and those for the southerly cave sites become climatologically more reasonable. For example, values of ΔT for the Iowa (-5 to -7 °C), Kentucky (-3 to -8 °C), and Texas (-11 to -16 °C) sites are in close agreement with those derived for the Great Lakes (-7 to -11 °C), Interior Lowlands (~-13 °C), and southwestern v.s. (-11 to -13 °C) regions from the ice-age climate modelling of Gates¹⁵ based on the CLIMAP²⁹ data, and not substantially different from those for the Great Lakes (-6 to -14 °C) and Appalachian Regions (8 °C) estimated by Peterson *et al.*³⁰ from pollen data. Of the six data sets in Table 2, that for Iowa is the most important because the Coldwater Cave stalagmite not only has the highest growth rate and smallest age uncertainties of any of the samples analysed, but also was deposited during the period of the change from the full glacial conditions of the late Wisconsinan to the interglacial conditions of the Holocene. The well-defined trend towards lower isotopic temperatures during the late Wisconsinan glaciation and the virtual coincidence of the speleothem temperature minimum and the glacial maximum at 18,000 yr BP give us confidence in these results. However, the data for the older speleothems must be considered inherently less accurate because of the larger uncertainties in dating. This especially applies to individual samples of 'intermediate-type' whose ages lie near climatic boundaries and thus may be misclassified as to climate type. The situation in Bermuda is different and more complex than in the continental sites and as a result somewhat problematic. Here, modern precipitation lies on a line of lower slope and d_0 value ($\delta D = 6\delta^{18}O + 4.2$, ref. 13). Because d_0 should in theory be zero

or positive depending on whether evaporation occurs totally in equilibrium or kinetic conditions¹, the glacial-age palaeotemperatures calculated for Bermuda in Table 2 are calculated for Bermuda in Table 2 assuming only a -4.2% shift in d_0 , while X is presumed to remain constant. Also, the Bermudian speleothems analysed were all formed at elevations below present sea level in presently flooded caves, so that none record fully interglacial conditions at this site.

The average difference between modern and ancient temperatures (ΔT) is negative and greatest for deposits formed during glacial times (Table 2). The apparently low ΔT values observed for some 'intermediate' samples presumably reflects the fact that we have set $d_0 = 0$ for these samples whereas during these times the true d_0 value was between 0 and +10. Without some estimate of both the $^{18}O/^{16}O$ and D/H ratios of glacial-age waters it is not possible to define the exact δD - $\delta^{18}O$ relationship. However, as long as the relationship remains linear with a slope $d\delta D/d\delta^{18}O = 8$, then small changes in the intercept d_0 will only have a comparatively small effect on the calculated temperature difference, ΔT . Using a global model of oceanic evaporation and precipitation, Merlivat and Jouzel¹ find that d_0 is virtually independent of the temperature at the site of evaporation, but decreases linearly with increasing relative humidity at the source site. A value of $d_0 = 0$ would correspond to a relative humidity of subtropical oceanic air of about 100% compared with a value of 81% at present. Growth of continental ice sheets might also have had an effect on the δD - $\delta^{18}O$ relationship due to the consequent enrichment of both D and ^{18}O in seawater.

It would thus seem important to obtain independent measurements of both $^{18}O/^{16}O$ and D/H ratios in dated Pleistocene palaeowaters, to evaluate the possibility of secular changes in the δD - $\delta^{18}O$ relationship for meteoric waters. Such work is in progress in the McMaster laboratory. Similar detailed studies on other ice cores are also required.

This research was funded by grants from the Climate Dynamics Programme of the NSF (ATM77-20365) and by the National Research Council (Canada). Support from the National Speleological Society and Cave Research Foundation is also acknowledged. We thank L. F. Binney, D. White, M. Boska and J. Borthwick for laboratory assistance, S. Epstein, J. R. O'Neil and J. Gray in whose laboratories some of the D/H ratio measurements were made, and N. J. Shackleton, and T. C. Atkinson for critical comments. Manuscript typed by Mrs E. Yuill and illustrations prepared by J. Bennett.

Received 5 June 1980; accepted 22 January 1981.

- Merlivat, L. & Jouzel, J. *J. geophys. Res.* **84**, 5029 (1979).
- Craig, H. *Science* **133**, 1702 (1961).
- Dansgaard, W. *Tellus* **6**, 436 (1964).
- Friedman, I., Redfield, A. C., Schoen, B. & Harris, J. *Rev. Geophys.* **2**, 177 (1964).
- Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39 (1973).
- Hendy, C. L. *Geochim. cosmochim. Acta* **35**, 801 (1971).
- Duplessy, J. C., Labeyrie, J., Lalou, C. & Nguyen, H. *Nature* **226**, 631 (1970).
- Harmon, R. S., Thompson, P., Schwarcz, H. & Ford, D. C. *Quat. Res.* **9**, 54 (1978).
- Schwarcz, H. P., Harmon, R. S., Thompson, P. & Ford, D. C. *Geochim. cosmochim. Acta* **40**, 657 (1976).
- O'Neil, J. R., Clayton, R. N. & Mayeda, T. K. *J. chem. Phys.* **30**, 5547 (1969).
- O'Neil, J. R., Adami, L. H. & Epstein, S. *J. Res. U.S. geol. Surv.* **3**, 623 (1975).
- Harmon, R. S. *Wat. Resour. Res.* **15**, 476 (1979).
- Duplessy, J. C., Labeyrie, J., Lalou, C. & Nguyen, H. V. *Quat. Res.* **1**, 162 (1971).
- Harmon, R. S., Schwarcz, H. P. & O'Neil, J. R. *Earth planet. Sci. Lett.* **42**, 254 (1979).
- Gates, L. G. *Science* **191**, 1138 (1976).
- Harmon, R. S., Thompson, P., Schwarcz, H. P. & Ford, D. C. *Nat. Speleol. Soc. Bull.* **37**, 21-33 (1975).
- Allen, C. R., O'Brien, R. & Sheppard, S. M. F. *Arctic Alpine Res.* **8**, 297 (1976).
- Ford, D. C., Harmon, R. S., Schwarcz, H. P., Wigley, T. M. L. & Thompson, P. *Arctic Alpine Res.* **16**, 219 (1976).
- Epstein, S., Sharp, R. P. & Gow, A. J. *Science* **142**, 1570 (1970).
- Epstein, S., Sharp, R. P. & Gow, A. J. *J. geophys. Res.* **70**, 1809 (1965).
- Stewart, M. K. N. Z. *J. geol. Geophys.* **18**, 49 ().
- Schoell, M. & Faber, E. *Geol. Jb.* **D17**, 197 (1976).
- Job, C., Moser, H., Rauter, W., Stichler, W. & Zotti, J. *Naturwissenschaften* **62**, 136 (1976).
- Gonfiantini, R., Conrad, G., Fontes, J. Ch., Sauzay, G. & Payne, B. R. *Isotope Techniques in Groundwater Hydrology* Vol. 1, 227 (IAEA, Vienna 1974).
- Epstein, S. & Mayeda, T. *Geochim. cosmochim. Acta* **4**, 213 (1953).
- Friedman, I. *Geochim. cosmochim. Acta* **4**, 89 (1953).
- Taylor, H. P. *Econ. Geol.* **69**, 843 (1974).
- Thompson, P., Schwarcz, H. P. & Ford, D. C. *Bull. geol. Soc. Am.* **87**, 1730 (1976).
- CLIMAP Project Members. *Science* **191**, 1131 (1976).
- Peterson, G. M. *et al. Quat. Res.* **12**, 47 (1979).

Two orogenies in the Grenville Belt?

A. J. Baer

Department of Geology, University of Ottawa, Ottawa, Canada K1N 6N5

The age of the Grenvillian event is commonly taken to be ~1,000 Myr, but depending on the isotopic systems used and the rocks examined, dates vary from 1,100 Myr or more¹ to ~700 Myr (ref. 2). K-Ar mineral dates and ⁴⁰Ar/³⁹Ar dates reflect blocking temperatures that correspond to regional cooling. They should, therefore, be younger than the true age of a thermal event that would be more accurately dated by Rb-Sr whole-rock isochrons or by U-Pb dating of zircons. However, because blocking temperature for argon is estimated to be 550–600 °C in hornblende³, hornblende dates should at least correspond to the time at which medium to high-grade metamorphism was last reached in the rock if systems have remained closed ever since. In the Grenville Province, the discrepancy between ~1,100 and ~950 Myr dates recorded by different radiometric clocks has commonly been attributed to a progressive and slow cooling of the region⁴. I suggest here that the presence of at least two distinct thermal events in that time interval is a plausible explanation.

It has been shown both experimentally⁵ and theoretically⁶ that different minerals have different blocking temperatures for Ar and Sr, and this property has been used to unravel the evolution of complex polymetamorphic crystalline terrains⁷. In the western Grenville Province, numerous such cases have been documented of K-Ar mineral ages at ~1,000 Myr BP whereas Rb-Sr whole-rock isochrons ran from 1,400 to 2,350 Myr BP (refs 8, 9). At least two distinct thermal events (Hudsonian and Grenvillian) have been thought to affect the region. There are other cases, however, and more particularly in the central part of the Grenville Province, where 40% of U-Pb dates on zircons and Rb-Sr whole-rock isochrons cluster around 1,050–1,100 Myr whereas 18% coincide with the peak of K-Ar dates at 900–950 Myr. A similar difference of ~150 Myr is also found in some K-Ar determinations of hornblende-biotite pairs from the same area (Table 1). Is this due to slow cooling?

Two peaks of K-Ar dates have been found^{10,11} from the Grenville Province and there are two possible interpretations. At the time, the lack of Rb-Sr whole-rock isochrons and U-Pb zircon dates corresponding to the younger K-Ar peak seemed to favour a slow cooling interpretation. Since then most authors have assumed that isostatic re-equilibration and erosion following orogenic activity could explain a 150 Myr time-lag between Rb-Sr whole-rock dates and K-Ar dates.

The determination of which length of time justifiably represents slow cooling, and which one separates two events is difficult. Blocking temperatures in metamorphic terrain are controlled by uplift of large crustal blocks⁴. In the Grenville, from metamorphic mineral assemblages the total crustal thickness during the Grenvillian event must have reached 60–70 km (ref. 12). By isostatic adjustment this must have caused the formation of an elevated belt as high as the Tibetan Plateau. Speed of erosion of this belt would control its rate of cooling. Calculations of recent cooling rates in the Alps give values varying from 50 °C Myr⁻¹, 29–26 Myr ago to 6 °C Myr⁻¹ for the past 14 Myr (ref. 13). Estimates based on rates of erosion and on an assumed geothermal gradient of 40 °C km⁻¹ yield values of 32–16 °C Myr⁻¹ at present¹⁴. Because cooling curves seem to be strongly exponential over a few tens of million years these values are certainly minimum figures. In the Grenville Province, dating of hornblende and biotite pairs by the ³⁹Ar/⁴⁰Ar technique indicates ~30–50 Myr between blocking temperatures of the two minerals^{3,15,16}, corresponding therefore to a cooling rate of

6–10 °C Myr⁻¹, in reasonable agreement with alpine estimates, if somewhat on the low side. If blocking temperature for argon in hornblende is of the order of 550–600 °C, the cooling of medium- to high-grade metamorphic rocks to ~300 °C should commonly take 50 Myr. However, given the accuracy of most available K-Ar dates, this time gap is too small to be statistically significant. Specific cases have been reported, however, where discrepancy between various dates cannot be explained by differences in cooling rates. For instance, in Lac Croche area Rb-Sr whole-rock isochrons show the peak of regional metamorphism and deformation to date from ~1,100 Myr ago (1,100 ± 27; 1,070 ± 43) whereas formation of a younger (metasomatic?) granite was dated from 940 ± 20 Myr by the same method¹. K-Ar dates on micas in gneisses give 955 Myr. In three other areas, hundreds of kilometres apart, in Cardiff and Mont Laurier in the Central Gneiss Belt and in Johan Baetz area in the eastern Grenville Province, uranium-bearing pegmatites have been dated from ~950 Myr ago (Rb-Sr whole-rock isochron method) whereas the peak of metamorphism dates from 1,100 to 1,050 Myr ago (same dating method applied to surrounding granites)¹⁷. Fowler and Doig¹⁷ report that uraniferous granites "seem to have crystallized from melts produced after the main pulse of the Grenville orogeny rather than through processes of *in situ* anatexis or metasomatism". In other words, geologically younger granitic rocks are isotopically distinct from and ~150 Myr younger than adjacent rocks. Finally, recent dating of the Marcy anorthosite in the Adirondacks gives 1,190 ± 130 Myr by the Sm-Nd method and 930 ± 40 Myr by the Rb-Sr whole-rock isochron method on the same samples¹⁸.

How general is the distribution of these two temporally distinct events? K-Ar dates in the 900–950 Myr range are distributed throughout the Province, but 1,050–1,100 Myr Rb-Sr whole-rock isochrons and U-Pb zircon dates are concentrated in a northeasterly trending swath in the central part of the Province only (Fig. 2). West of a line running from Bancroft to Chibougamau, Rb-Sr whole-rock isochrons reflect still older thermal events extending back in time to the Archean^{8,9,19–22}. Thus, in the western Grenville Province the absence of a 1,100 Myr set of dates apparently means that the 950 Myr dates stand on their own, and do not represent the end of a long cooling phase. Previous thermal events of high intensity last affected the area around 1,300 Myr or earlier⁹.

Outside Canada, two distinct radiometric events have been reported from the Adirondacks (1,100 and 850–900 Myr)^{23,24}, from east Greenland (1,150 and 950 Myr)^{25,26}, from southern Norway (1,100 and 900–100 Myr)^{27,28} and from the basement of the Appalachians (1,150 and 835 Myr)^{11,29}. Dates from Skåne in

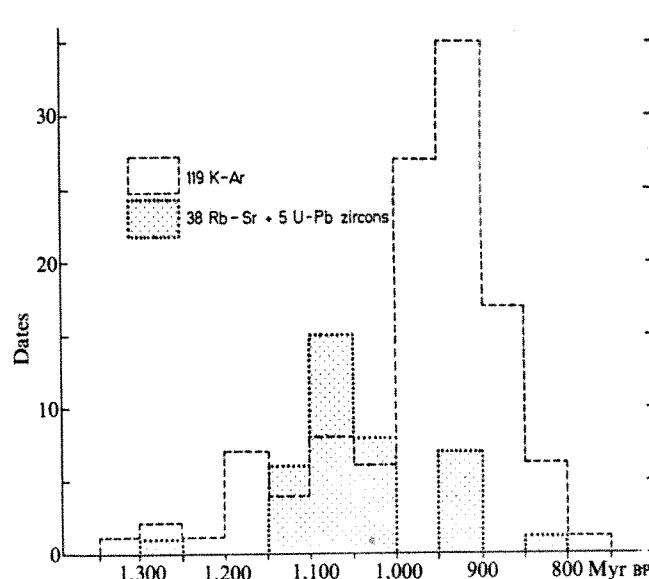


Fig. 1 Histogram of isotopic ages in the Grenville Province.

Table 1 Comparison of isotopic ages from some areas inside and outside the Grenville Province

Outside the Grenville Belt					
	U-Pb (Myr)	Rb-Sr whole-rock isochron (Myr)	K-Ar mica (Myr)	K-Ar amphibole (Myr)	Maximum difference (Myr)
Basler Lake		2,510 ± 106 ³⁷	2,460, 2,420, 2,390 ³⁷		120
Nueltin Lake		1,775 ± 59 ³⁷	1,800, 1,735 ^{32,38}		25
Mistastin Lake		1,318 ± 15 ³⁹	1,340 ⁴⁰	1,325 ⁴⁰	17
Umiakovik Lake		1,290 ± 30 ⁴¹	1,275 ⁴²		15
Inside the Grenville Belt					
North Bay	1,160 ⁴³		940 ⁴²	884 ⁴⁴	276
French River	1,100		985, 930 ⁴⁵		170
	1,160 ⁴³				
Manhattan Prong	1,150 ¹¹		835 ⁴⁶		315
Lac Croche		1,119 ± 24 ⁴⁷	955 ⁴⁰		164
Deloro pluton		1,096 ± 46 ³⁷		989 ¹⁰ , 875 ⁴⁰	221
Labrieville		1,165 ± 90 ⁴⁸	875 ⁴²		290
Reading Prong	1,060 ⁴⁹		788 ⁴⁶	873 ⁴⁶	272
New Jersey Highlands	1,170 ¹¹		905 ¹¹		265

All Rb-Sr dates recalculated for $\lambda = 1.42 \times 10^{-11} \text{ yr}^{-1}$.

Sweden suggest similar relationships, but results are somewhat ambiguous³⁰. This remarkable persistence of two peaks of thermal activity throws some doubt on an interpretation involving slow uplift and cooling and it suggests that the possible existence of two thermotectonic events should be carefully evaluated. This would mean that an early Grenvillian event would only have affected some parts of the Grenville Belt, and a later event, of lesser intensity but greater areal extent would be responsible for the later group of dates.

In Scandinavia regional metamorphism and deformation peaked around 1,200 Myr ago, and again around 1,000 Myr ago²⁸. According to Berthelsen (personal communication) early Grenvillian structures (Ostfoldian) trend north, with a westerly vergence, and later structures cut across them. Structural trends are similar in both cases but the younger thrust planes dip west instead of east. In Canada, early Grenvillian Rb-Sr isochron dates fall in a zone of north-northeasterly structures that run obliquely to the Grenville Front (Fig. 2). This penetrative structural grain contrasts with a variety of trends in other parts of the Province and suggests that the earlier event was more homogeneous and more intense than the later one. The latter did not generally open Rb-Sr whole-rock systems, and in places it did not even reset K-Ar mineral dates (1,533 ± 42 Myr on hornblende³¹ and 1,395 Myr on biotite³² for example).

If the distribution of 'Grenvillian' dates corresponds to two events 150 Myr apart, thermochrons⁴ drawn from the distribution of all K-Ar biotite dates are meaningless. In fact, one cannot draw continuous thermochrons for either a 1,100–1,150 Myr group or a 850–1,000 Myr group of K-Ar dates on biotites.

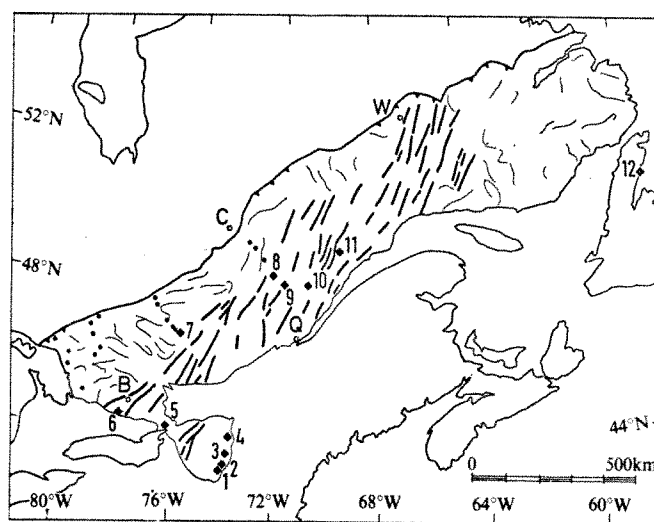


Fig. 2 Distribution of 1,100 ± 50 Myr Rb-Sr whole-rock isochron dates in the Grenville Province. Dots correspond to isochrons giving older dates. Note coincidence of north-northeasterly structural trends with rocks dated at 1,100 ± 50 Myr dates. B, Bancroft; C, Chibougamau; Q, Quebec City; W, Wabush. 1♦, 1,120 Myr; 2♦, 1,095 ± 39 Myr; 3♦, 1,130 ± 10 Myr; 4♦, 1,100 ± 15 Myr; 5♦, 1,073 ± 50 Myr; 6♦, 1,080 ± 39 Myr; 7♦, 1,131 ± 56 Myr; 8♦, 1,159 ± 99 Myr; 9♦, 1,021 ± 10 Myr; 10♦, 1,057 ± 27 Myr; 11♦, 1,165 ± 90 Myr; 12♦, 1,106 ± 90 Myr.

Table 2 a, Concordant dates from different isotopic systems in various parts of the Grenville Belt. b, Discordant K-Ar dates from mineral pairs in the same rock; Grenville Front zone

a			b				
	Rb-Sr whole-rock isochron (Myr)	K-Ar mica (Myr)	K-Ar amphibole (Myr)	Muscovite (Myr)	Biotite (Myr)	Amphibole (Myr)	Maximum difference (Myr)
Lac Croche, P.Q.	961 ± 20 ¹	955 ⁵⁰		Killarney	1,180 ⁴⁴	1,170 ³² 965 ⁴⁴	215
St. André, P.Q.	927 ± 15 ⁵¹	940 ³⁸		Chibougamau		1,105 ⁴²	184
Mattawa, Ont.	1,029 ± 152 [*]	940 ⁴²	965 ⁴⁴	Gagnon		1,051 ⁴⁴	220
Scoresby Sund	1,169 ± 6 ²⁶		1,124 ± 25 ²⁵	Wabush	980 ³⁸	1,125 ³⁸	145

* R. K. Wanless (personal communication).

The available geochronological data suggest therefore, to a first approximation, that the Grenvillian thermal activity extended from ~1,150 to 850 Myr ago. Considerations of argon retention by various minerals and of the significance of Rb–Sr whole-rock isochron and U–Pb zircon dates suggest either that metamorphism peaked around 1,100 Myr ago and that all younger dates represent a tail-end of this same event, or that two distinct metamorphic events took place, a stronger one around 1,100 Myr ago and a weaker one around 900 Myr ago. If 'slow cooling' is the explanation and is taken to mean a cooling rate of ~300 °C between 1,150 and 950 Myr ago, this is one order of magnitude slower than present cooling rates in the Alps. If cooling is in turn dependent on rates of isostatic uplift and of erosion, slow uplift and slow erosion of a very thick continental crust seem to contradict what we know of crustal isostasy. The other possible explanation is that of two distinct events. This would explain why some Rb–Sr whole-rock isochron dates and U–Pb zircon dates are identical to K–Ar biotite dates whereas others differ by 200 Myr (Table 2). In this interpretation, the younger event would have affected all the belt but may not have been strictly coeval everywhere, because dates reported from the southern Appalachians (700–750 Myr)² are considerably younger than those from the northern Appalachians^{11,29,33}, Canada^{34,35} and Scandinavia (800–900 Myr)^{28,30,36}. The older events would have been of more restricted extent but, at least in the central Grenville Province, deformation was intense enough to give a common structural grain to the region (Fig. 2). Figure 1 summarizes the argument. Rb–Sr whole-rock and U–Pb dates on zircon form two peaks. The older is higher because dates were generally not reset by the second event. K–Ar dates form one peak that coincides with the younger of the previous two. The K–Ar histogram is asymmetric, suggesting the existence of a set of older dates partly reset by the younger event.

I thank K. Bell, G. W. Berger, R. D. Dallmeyer and W. K. Fyson for constructive comments, E. Hearn for drawing the figures, S. Meunier and H. DeGouffe for typing the text.

Received 30 June; accepted 30 December 1980.

- Barton, J. M. Jr & Doig, R. *Can. J. Earth Sci.* **9**, 1180–1186 (1972).
- Dallmeyer, R. D. *Bull. geol. Soc. Am.* **86**, 1740–1744 (1975).
- Dallmeyer, R. D. & Sutter, J. F. *J. geophys. Res.* **85**, B6, 3177–3186 (1980).
- Harper, C. T. *Earth planet. Sci. Lett.* **3**, 128–132 (1967).
- Hart, S. R. *J. Geol.* **72**, 493–525 (1964).
- Dodson, M. H. *Contr. Miner. Petrol.* **40**, 259–274 (1973).
- Armstrong, R. L. in *Potassium–Argon Dating* (eds Schaeffer, O. A. & Zähringer, J.) 117–131 (Springer, Heidelberg, 1966).
- Grant, J. A. *Science* **146**, 1049–1053 (1964).
- Davis, G. L., Hart, S. R. & Aldrich, L. T. *Yb. Carnegie Instn Wash.* **65**, 379–386 (1967).
- Macintyre, R. M., York, D. & Moorhouse, W. W. *Can. J. Earth Sci.* **4**, 815–828 (1967).
- Long, L. E. & Kulp, J. L. *Bull. geol. Soc. Am.* **73**, 969–996 (1962).
- Grandth, J. W. & Barstow, N. *Bull. geol. Soc. Am.* **91**, 75–77 (1980).
- Miller, D. S., Wagner, G. A. & Jäger, E. *U.S. Geol. Surv. Open File Rep.* **78-701**, 297–299 (1978).
- Clark, S. P. Jr in *Lectures in Isotope Geology* (eds Jäger, E. & Hunziker, J. C.) 225–230 (Springer, Berlin, 1979).
- Berger, G. W., York, D., Dunlop, D. J. & Buchan, K. L. *U.S. Geol. Surv. Open File Rep.* **78-701**, 30–32 (1978).
- Berger, G. W., York, D. & Dunlop, D. J. *Nature* **277**, 46–48 (1979).
- Fowler, A. D. & Doig, R. *Geol. Ass. Can. Prog. Abstr.* **4**, 50 (1979).
- Ashwal, L. D., Wooden, J. L. & Shih, C.-Y. *Geol. Soc. Am. Abstr. Prog.* **12**, 380 (1980).
- Krogh, T. E., Davis, G. L., Aldrich, L. T. & Hart, S. R. *Yb. Carnegie Instn Wash.* **66**, 528–536 (1968).
- Krogh, T. E. & Davis, G. L. *Yb. Carnegie Instn Wash.* **67**, 224–230 (1969).
- Davis, G. L., Krogh, T. E. & Hart, S. R. *Yb. Carnegie Instn Wash.* **68**, 307–314 (1970).
- Krogh, T. E., Davis, G. L. & Hart, S. R. *Yb. Carnegie Instn Wash.* **69**, 337–344 (1971).
- Silver, L. T. *New York State Museum Service, mem.* **18**, 233–251 (1969).
- Gaudette, H. E., Michard-Vitrac, A. & Allègre, C. J. *Geol. Soc. Am. Abstr. Prog.* **11**, 431 (1979).
- Hansen, B. T., Oberli, F. & Steiger, R. H. *Geol. Surv. Greenland Rep.* **58**, 55–58 (1973).
- Hansen, B. T., Oberli, F. & Steiger, R. H. *Geol. Surv. Greenland Rep.* **66**, 32–38 (1974).
- Pedersen, S. et al. *U.S. Geol. Surv. Open File Rep.* **78-701**, 329–331 (1978).
- Versteve, A. J. *Norg. geol. Unders.* **318**, 1–50 (1975).
- Dallmeyer, R. D., Sutter, J. F. & Baker, D. S. *Bull. geol. Soc. Am.* **86**, 1435–1443 (1975).
- Klingspor, I. *Geol. Förr. Stockh. Förrh.* **98**, 195–216 (1976).
- Emlie, R. F. *Can. J. Earth Sci.* **15**, 438–453 (1978).
- Lowdon, J. A., Stockwell, C. M., Tipper, H. W. & Wanless, R. K. *Geol. Surv. Can. Paper* **62-17** (1963).
- Dallmeyer, R. D. *Can. J. Earth Sci.* **15**, 1374–1379 (1978).
- Wanless, R. K., Stevens, R. D., Lachance, G. R. & Delabio, R. N. *Geol. Surv. Can. Paper* **71-2** (1972).
- Wanless, R. K., Stevens, R. D., Lachance, G. R. & Delabio, R. N. *Geol. Survey Can. Paper* **73-2** (1973).
- O'Nions, R. K. & Baadsgaard, H. *Contr. Miner. Petrol.* **34**, 1–21 (1971).
- Wanless, R. K. & Loveridge, W. D. *Geol. Surv. Can. Paper* **72-23** (1972).
- Lowdon, J. A. *Geol. Surv. Can. Paper* **61-17** (1961).
- Marchand, M. *Geol. Ass. Can. A. Meet. Prog. abstr.*, 58 (1974).
- Wanless, R. K., Stevens, R. D., Lachance, G. R. & Rimsaite, J. Y. H. *Geol. Surv. Can. Paper* **64-17** part 1 (1965).
- Wanless, R. K. *Geol. Surv. Can. Paper* **77-14** (1978).
- Leech, G. B., Lowdon, J. A., Stockwell, C. H. & Wanless, R. K. *Geol. Surv. Can. Paper* **63-17** (1963).
- Tilton, G. R. & Grünfelder, M. H. *Science* **159**, 1458–1461 (1968).
- Wanless, R. K., et al. *Geol. Surv. Can. Paper* **74-2** (1974).
- Lowdon, J. *Geol. Surv. Can. Paper* **60-17** (1960).
- Dallmeyer, R. D. et al. *Bull. geol. Soc. Am.* **86**, 1435–1443 (1975).
- Barton, J. M. Jr & Doig, R. *Can. J. Earth Sci.* **9**, 1180–1186 (1972).
- Bron, F., Giraud, P., Laurier, A. & Violette, C. *Rebd. Séanc. Acad. Sci. Paris* **276**, 689–692 (1973).
- Tilton, G. R. et al. *J. geophys. Res.* **65**, 4173–4179 (1980).
- Wanless, R. K. et al. *Geol. Surv. Can. Paper* **67-2** part A (1968).
- Frith, R. A. thesis McGill Univ. Montreal (1971).

A Pleistocene lacustrine episode in southeastern Libya

C. Gaven, C. Hillaire-Marcel & N. Petit-Maire

Centre National de la Recherche Scientifique, Laboratoire de Géologie du Quaternaire, Faculté des Sciences de Luminy-Case 907, 13288 Marseille, Cedex 2, France

Fossiliferous lacustrine deposits of middle Pleistocene age have been found in Central Fezzan (Libya)¹ on each side of the east–west depressions of the Wadi Shati, south of the Pre-cambrian Qarqaf anticline^{1–8}. They represent the shoreline of a former brackish water lake (Fig. 1) which, some 130,000 yr BP, covered an area about 100 km long and 10–20 km wide. At present, the area may be classified as hyperarid. Since 1978, a collaborative study of the fossil assemblages has been undertaken. It included macro- and micropaleontological determination, Th/U dating, stable isotope and chemical analyses, in parallel with sedimentological and archaeological studies of the sites. Here we consider the Th/U dating and the isotopic palaeohydrology of the lacustrine episode.

Two main facies have been observed. The first and most extensive occurs in the main central and western depressions of the Shati. It consists of isolated outcrops, 2–8 m in height and 5–15 m in diameter. They are almost totally composed of layers of mollusc shells, cross-shells, broken or whole shells, sometimes with both valves joined, with a variable but small amount of quartz sand. Typical current orientations of pelecypod valves are observed, which, in conjunction with the abundance of fragments, indicate a littoral environment. These deposits occur at elevations 330–347 m, some 40–60 m above the bottom of the Shati depression. They represent a single major lacustrine episode with large lake-level fluctuations (~10–15 m), as suggested by the alternance of topsets and bottomsets.

The second facies occurs in two small and partly isolated depressions at the eastern margin of the Shati Valley. It consists of quartz sands containing small pebbles which originate from the Nubian sandstones (Lower or Mid-Cretaceous continental deposits). Scattered shells or small layers of shells are present here and there. The easternmost depression reveals freshwater molluscs; the intermediate one shows both fresh- and brackish-water fauna. Fluvial inputs characterize this margin of the lake, where freshwater fauna developed in small embayments. Note that these depressions could have been partly filled with water during several episodes after the main lacustrine phase.

In the main lake, the fauna includes molluscs³, ostracods⁵ and foraminifera⁵. Among the molluscs, *Cerastoderma glaucum*, the Mediterranean cardium, is by far the most abundant species. It is usually associated with *Melania tuberculata* and *Hydrobia* sp. This fossil assemblage reflects a shallow water community (depth <10 m) living on a sandy substrate³. It also suggests a temperature of ~20 °C (which is implied for larval development of cardium) and a minimum salinity³ of 3‰. In contrast, the deposits in the small depressions of the eastern margin are

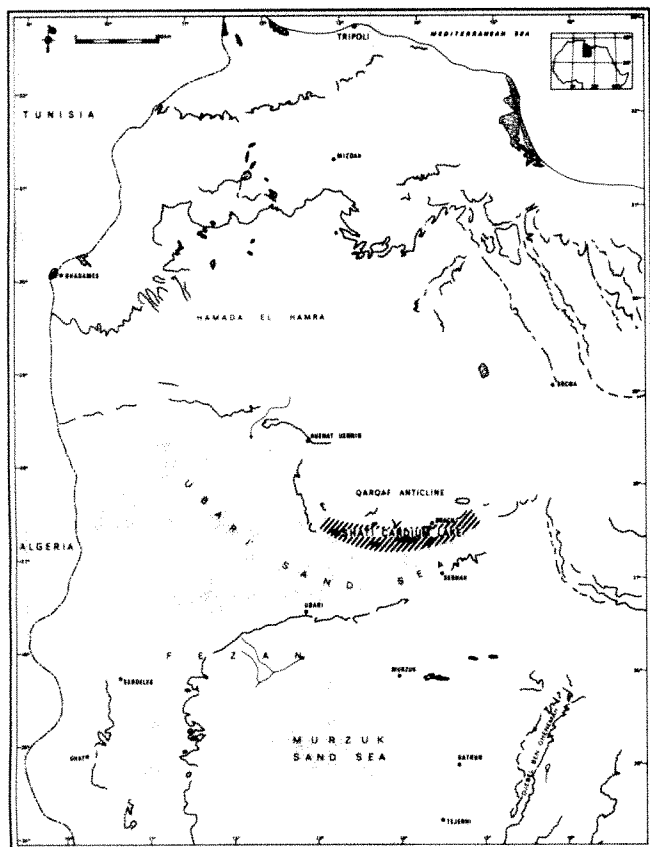


Fig. 1 Proposed area for the Shati Cardium Lake, .

characterized by freshwater assemblages which include *Corbicula africana*, *Coelatura lacoini*, *Bulinus truncatus* and *Limnaea natalensis*.

The shell morphology of *Cyprideis*, which is the only occurring genus of ostracods, indicates a strongly variable physical environment. The salinity ranged from 5 to 10‰ during the periods of *Cyprideis* development^{8,9}.

Surprisingly, five species of marine foraminifera have been identified in these lacustrine sediments which are located almost

1,000 km from the sea: *Ammonia beccarii tepida*, *Protelphidium paraliu*, *Criboelphidium articulatum*, *Discorbis* sp. and *Rosalina globularis*⁵. Such occurrences, so far inland, of marine organisms in saline lakes are probably due to transportation by migrating birds. From a palaeoecological point of view, the Shati Cardium Lake fossil assemblages reflect brackish but highly variable conditions. Charophytes (*Lamprothamnium* sp.) also occur at a few localities^{10,11}; they could correspond to a 5–10‰ salinity range¹² in agreement with the ostracod requirements.

Stable carbon and oxygen isotope analyses have been made on 46 samples of fossil shells and calcitic cement. For each sample, Mg^{2+} , Sr^{2+} contents and calcite/aragonite ratios have been determined. As Fig. 2 shows, some shells present high $\delta_{PDB}^{18}O$ and $\delta_{PDB}^{13}C$ values that are almost characteristic of 'marine' conditions. The Cardium Lake seems to be a perfect example of isotopic fractionation in a closed basin, where evaporation induces an increase of ^{18}O and ^{13}C content in the water phase¹⁴. A direct $\delta^{18}O$ salinity relationship cannot be established as the original ionic concentrations are unknown. The samples from the eastern part of the lake show typical freshwater conditions with lower $\delta^{18}O$ and ^{13}C values. Therefore, one may assume that the strongest freshwater inputs came from the east. In general, isotope ratios of the shells reflect brackish but highly variable conditions. Many samples present clear evidence of recrystallization. They are characterized by low $\delta^{18}O$ values and high ^{13}C content, almost identical to the values which are observed in the calcite which cements the shells. Calcite/aragonite ratios systematically confirm this interpretation. Moreover, lower Sr^{2+} and higher Mg^{2+} content may be observed in these recrystallized shells. The isotopic shift which occurs during aragonite–calcite transformation could thus determine the stability or originally calcitic shells. Hence, it is possible to use such calcitic shells for Th/U dating, as the ^{18}O and ^{13}C content reflects the mineral stability and the risk of contamination. In the Cardium Lake fossils, the large ^{18}O variations due to evaporation totally erase the temperature effect, the amplitude of which is much less.

More than 20 Th/U dates have been obtained on the cardium shells, by the ionium disequilibrium method¹⁵, with systematic X-ray control of calcite occurrence in the originally aragonitic shells. Most of the dated samples presented a negligible percentage of calcite—only occasionally does the calcite/aragonite ratio exceed 0.1. But even then, the ages seem to

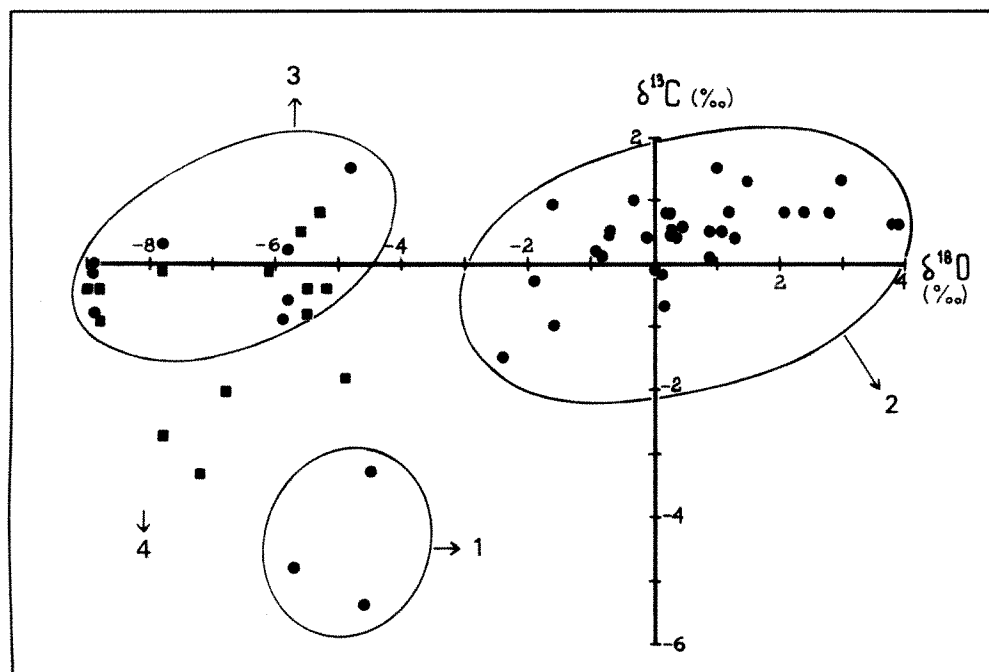


Fig. 2 Carbon and oxygen isotopes of the Shati Cardium Lake fossils. Three groups are indicated: 1, freshwater conditions; 2, brackish water conditions; and 3, fossils with aragonite–calcite transformation. ●, Fossils; ■, calcitic cement.

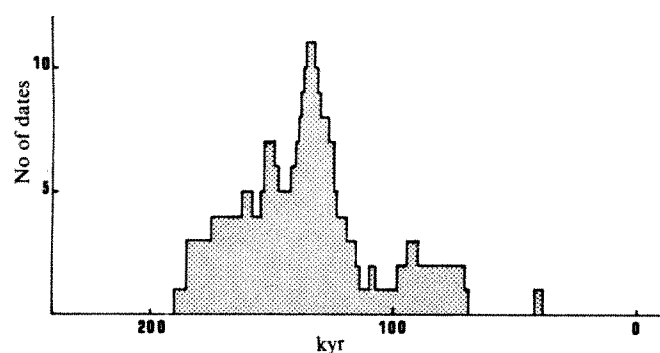


Fig. 3 Frequency histogram of U/Th dates $\pm 1\sigma$. Total number of dates, 21.

agree with those obtained on purely aragonitic shells. This would indicate a very precocious (almost syngedimentary) phase of aragonite-calcite transformation. Moreover, the ^{238}U content (<1 p.p.m.) between samples is comparable¹⁶. The $^{234}\text{U}/^{238}\text{U}$ ratios are greater than 1.15 and clearly reflect the continental nature of the water⁶. The $^{232}\text{Th}/^{230}\text{Th}$ ratios are very small (few p.p.m.), indicating negligible Th contamination. Figure 3 shows that the major lacustrine episode took place some 130 kyr BP. Although it is difficult to be precise about the duration of this phase, it clearly preceded the 125 kyr climatic optimum of the last (Eemian) interglacial. It seems more realistic to associate it to the transition between Stages 6 and 5e of the oceanic ^{18}O record¹⁷. A few dates averaging around 90 kyr and around 40 kyr suggest two more recent but less extensive humid episodes, with small lakes within some of the depressions.

As in most Saharan sites, the prehistoric artefacts are surface finds, for the greatest part, but in some cases they have been observed *in situ* in the lake deposits. They belong to at least two different cultures. Primarily there are biface tools of middle Pleistocene age, presenting typical late Acheulean features. Second, characteristic Levallois-Mousterian lithics are also present. These archaeological periods seem to fit the lake chronology.

From a palaeoclimatic point of view, the major phase of the Shati Cardium Lake developed at the end of the middle Pleistocene and as evidenced by sedimentology, stable isotopes and faunal assemblages, was characterized by large fluctuations and considerable evaporation, leading to almost 'marine' conditions. It seems realistic to associate these fluctuations mainly with groundwater oscillations.

We thank Professor Rhodes W. Fairbridge for helpful comments.

Received 20 October 1980; accepted 9 January 1981.

- Petit-Maire, N. *Rapport Laboratoire Géologie Quaternaire*, Marseille (1977).
- Petit-Maire, N., Casta, L., Delibrias, G. & Gaven, C. *Proc. 2nd Symp. Geology of Libya* (1978).
- Gaillard, J. & Testud, A. M. *Proc. 2nd Symp. Geology of Libya* (1978).
- Icole, M., Petit-Maire, N. & Riser, J. *8th Réunion. a. Sci. Terre*, Marseille (Société Géologique Française, Paris, 1980).
- Blanc-Vernet, L., Carbonel, P., Gaillard, J., Petit-Maire, N. & Peypouquet, J. P. *8th Réunion. a. Sci. Terre*, Marseille (Société Géologique Française, Paris, 1980).
- Petit-Maire, N., Delibrias, G., Gaven, C. in *Sahara and the Surrounding Seas* (eds Sarnthein, M., Seibold, E. & Rogion, P.) 289–295 (Balkema, Rotterdam 1980).
- Hillaire-Marcel, Cl. *Bull. A.S.E.Q.U.A.* 56–57, 79–83 (1979).
- Petit-Maire, N. *et al. Abstr. 26th int. Geol. Congr.*, 2, 681 (1980).
- Peypouquet, J. P. thesis, Univ. Bordeaux (1977).
- Carbonel, P. & Pinson, J. *Proc. 7th int. Symp. on Ostracods* 211–217 (Serbian Geological Society, Belgrade, 1979).
- Soulié-Marsche, I. thesis, Univ. Montpellier (1979).
- Soulié-Marsche, I. in *The Shati Cardium Lake* (ed. Petit-Maire, N.) (in the press).
- Bourrouilh-Le Jan, F. C. r. heb. Séanc. Acad., Sci., Paris D287, 907–910 (1978).
- Degens, E. T., von Herzen, R. P., Wong, H. K., Deuser, W. O. & Jannasch, H. W. *Geol. Rdsch.* 62, 1–245 (1979).
- Cherdynstev, V. V. *Uranium 234 Atomizdat*. (IPST, Moscow, 1969).
- Rothe, P., Hoefs, J. & Sonne, V. *Sedimentology* 21, 373–395 (1974).
- Shackleton, N. J. & Opdyke, N. D. *Geol. Soc. Am. Mem.* 145, 449–464 (1975).

Microbiological and zooplankton activity at a front in Liverpool Bay

G. D. Floodgate, G. E. Fogg, D. A. Jones, K. Lochte & C. M. Turley

Department of Marine Biology, Marine Science Laboratories (University College of North Wales), Menai Bridge, Gwynedd LL59 5EH, UK

Shallow sea fronts, zones of transition between mixed and stratified waters, are generally characterized by accumulation of biomass^{1,2}. Various mechanisms can be suggested for this accumulation. Pingree *et al.*¹ considered that phytoplankton are limited on the stratified side by nutrient supply and on the mixed side by light intensity, whereas at the front itself, recently stabilized mixed water provides conditions for rapid growth. Adjacent water masses may also complement each other's nutrient deficiencies. At the western Irish Sea front for example, silica has been limiting on the mixed side and combined nitrogen on the stratified side, so that the front supported more algal growth than either water separately². Another possibility, as fronts are regions of convergence, is advection of suspended matter, which may include detritus, phytoplankton and zooplankton³. This is unlikely to end with the passive accumulation of material at the front: bacterial mineralization of organic detritus will stimulate phytoplankton growth, bacterial and phytoplankton growth will support increased zooplankton activity and this in turn will increase the rate of recycling. Although little investigated, this mechanism seems to be important in a front existing in Liverpool Bay, as we report here.

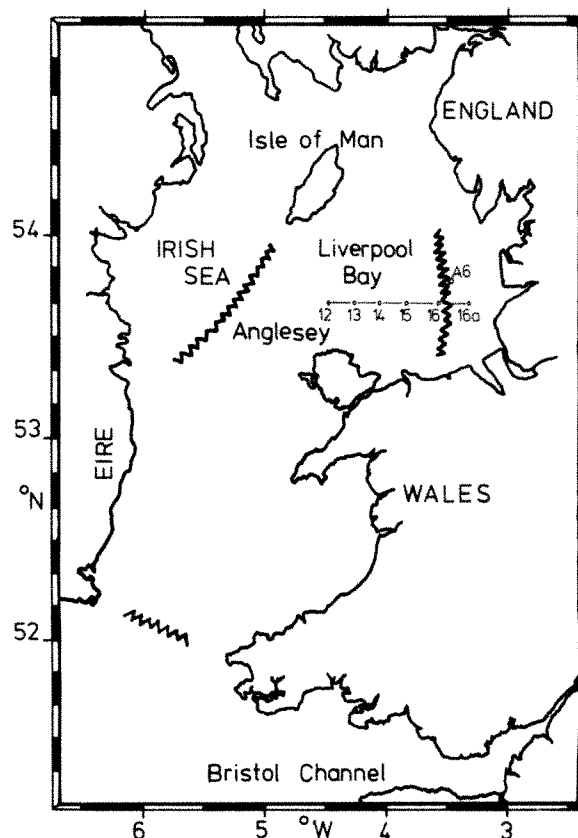


Fig. 1 Position of the Liverpool Bay and western and south-western Irish Sea fronts. Position of anchor station (A6) and station positions along reference line are indicated.

The general hydrographic and chemical features of Liverpool Bay have been described by Foster *et al.*⁴. A discontinuity or front, running approximately due north along the 4°W line of longitude from Great Orme Head, is a fairly regular feature in the Bay. In winter this front depends largely on salinity differences arising from river inflow whereas in summer, surface heating plays the major part in maintaining it. A study completed since the observations reported here shows that the situation is complex and that stratification associated with the front only becomes established in calm weather (S. Czitrom, unpublished results). The intermittent nature of this front undoubtedly complicates the biological picture and, for this reason, further investigation of bacterial and zooplankton activities at fronts is being conducted on the more stable western Irish Sea front (Fig. 1).

Observations were made along a transect at right angles to the Liverpool Bay front, which formed part of a grid of stations regularly occupied for hydrographical and phytoplankton surveys by R. V. Prince Madog (Fig. 1). Station positions were corrected to refer to the same tidal phase. Surface temperature and salinity were recorded continuously; at each station and half-way between regular stations, vertical profiles of temperature and salinity were determined (S. Czitrom, unpublished results). Microbiological samples were taken with sterile 2.5-l glass bottles attached to Johnson-Zobell microbiological samplers. Total bacterial numbers were estimated by an acridine orange epifluorescence direct counting technique⁵⁻⁷. Bacteria were classified into four size classes and volumes calculated by approximating to regular geometric shapes with ranges of volumes as follows: 'mini-bacteria' (cocci and rods) 0.004–0.1 μm^3 , cocci 0.1–0.4 μm^3 , rods 0.03–0.2 μm^3 and large rods

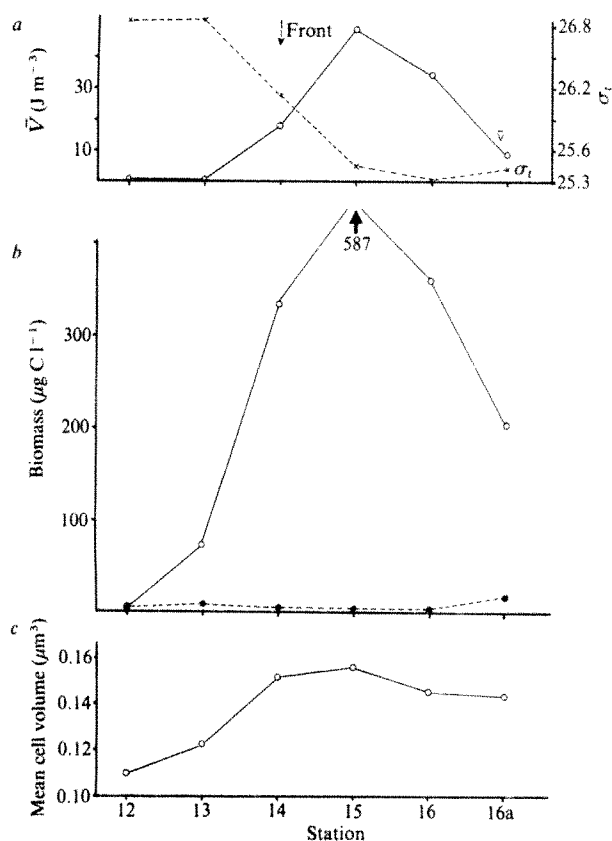


Fig. 2 Distribution of *a*, stability factor, $\bar{V}$ (○) and surface density, σ_t (×); *b*, bacterial (○) and phytoplankton (●) biomass in surface water; and *c*, bacterial mean cell volume (○) in surface water along the reference line indicated in Fig. 1 on 13–14 February 1978. $\bar{V}$ is the energy input required per unit depth to bring about complete vertical mixing of the stratified water column.

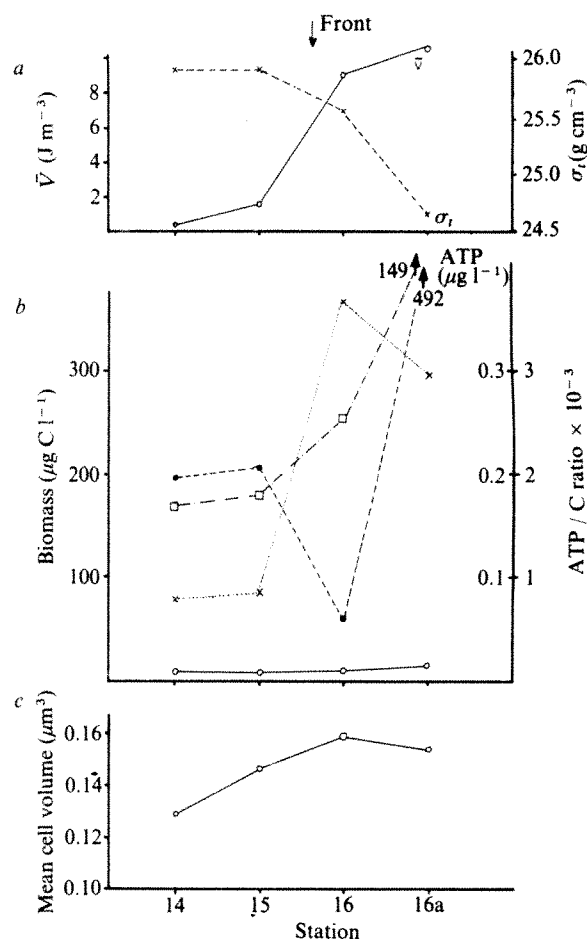


Fig. 3 Distribution of *a*, stability factor $\bar{V}$ (○, see Fig. 2 legend) and surface density, σ_t (×); *b*, bacterial (○) and phytoplankton (●) biomass, microbial cellular ATP (□) and cellular ATP/carbon ratios (×); and *c*, bacterial mean cell volume (○) in surface waters along the reference line indicated in Fig. 1 on 5 June 1978.

0.2–0.7 μm^3 . ATP content was determined according to Holm-Hansen and Booth⁸ and chlorophyll *a* content according to Yentsch and Menzel⁹. Bacterial biomass was calculated in terms of carbon from total bacterial volume assuming a density of 1.1 g cm^{-3} , a dry/wet ratio of 0.2 and a carbon/dry weight ratio of 0.5 (refs 10, 11). Phytoplankton biomass was calculated in terms of carbon using an organic carbon/chlorophyll *a* ratio of 30 (ref. 12). Zooplankton samples were obtained by surface tows and vertical hauls made through the whole water column.

In the late winter the vertical distribution of density (σ_t) along the transect showed mixed water offshore and stratified water inshore with the front near station 14, a situation clearly related to the high inflow of fresh water into the Bay during the winter. A similar picture of bacterial distribution, shown in Fig. 2, emerged from all four surveys made in winter/early spring. Bacterial biomass showed a peak on the inshore side of the front, the position of which was clearly defined by steep gradients in density and stability factor $\bar{V}$ (ref. 13). Bacterial biomass, in fact, correlates positively with $\bar{V}$ when stratification occurs. There was also a consistent difference in distribution of the size classes of bacteria, with the minibacteria predominating offshore, and the larger forms inshore, of the front. Bacterial biomass was always in considerable excess of phytoplankton biomass, which showed low levels without obvious relation to the front. In late spring/early summer the picture changed and in four surveys during this period was as shown in Fig. 3. Bacterial biomass was low, without any major difference on the two sides of the front. There was, however, the same pattern in size distribution as was

found earlier in the year. Phytoplankton biomass now greatly overshadowed bacterial biomass and was greatest adjacent to the front on the stratified side, a distribution similar to that found by Pingree *et al.*¹ at the front in the approaches to the English Channel. ATP concentration generally followed phytoplankton concentration but was higher than expected at the front itself, where there was a peak in microbial ATP/cellular carbon ratio.

When mean values for chlorophyll, bacterial numbers and ATP/cellular carbon ratio in the water column were plotted for the different stations as a function of time, the higher values for both chlorophyll and bacterial numbers were seen to have been always on the landward side of the transect. Bacterial numbers were highest in early spring and chlorophyll concentration was greatest in late spring and early summer with the ATP/cellular carbon ratio reaching a peak in between. The mean cell volume of the bacteria also increased with time to reach a peak at the same time as the chlorophyll.

Evidence that microbial activity as distinct from biomass was high at the front was provided by studies of urea utilization. These were made by determining release of $^{14}\text{CO}_2$ (measured as $^{14}\text{CO}_2$ released on acidification plus particulate ^{14}C) from ^{14}C -urea added to samples incubated on deck in the light (700 lx) and in the dark at 21 °C. Both phototrophic and heterotrophic breakdown of urea per unit volume of water, which were positively correlated, occurred at rates three to four times higher at the front than away from it (Fig. 4). Urea is metabolized by both bacteria and algae and evidence presented by Butler *et al.*¹⁴ indicates that urea, among other nitrogenous organic substances, may also be an important source of nitrogen for some phytoplankton species. Urea concentrations, measured by the method of McCarthy¹⁵, were highest at the front ($\sim 1 \mu\text{g l}^{-1}$ at 5 and 10 h in the series of observations shown in Fig. 4, compared with $\sim 0.1 \mu\text{g l}^{-1}$ at other times). Urea in offshore

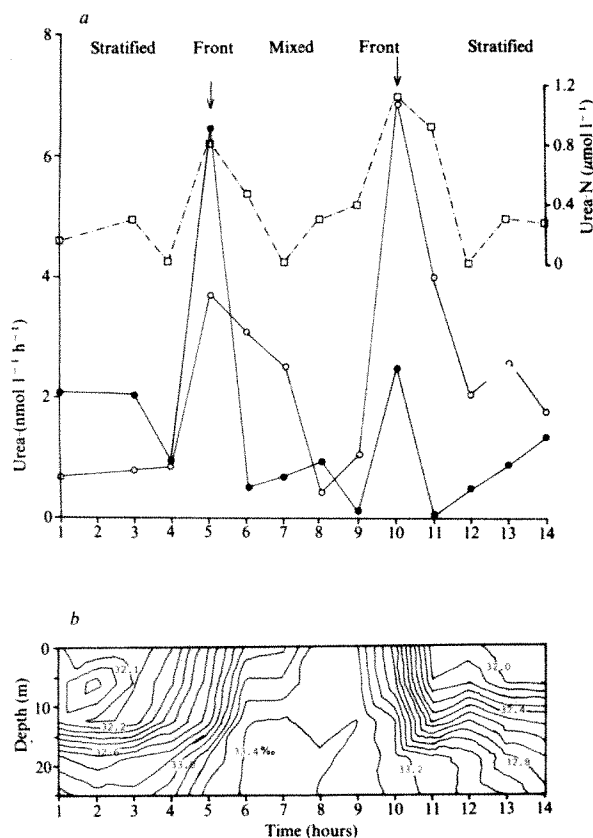


Fig. 4 Fourteen-hour long anchor station at station A6 (see Fig. 1) from 0612 to 1900 h, 4 May 1977, showing a front passing back and forth under tidal influence. *a*, Distribution of phototrophic (○) and heterotrophic (●) urea assimilation rates and *in situ* urea concentration (□) at 10 m depth; *b*, salinity profile: contours are at intervals of 0.1‰ and have been drawn by an objective contouring programme.

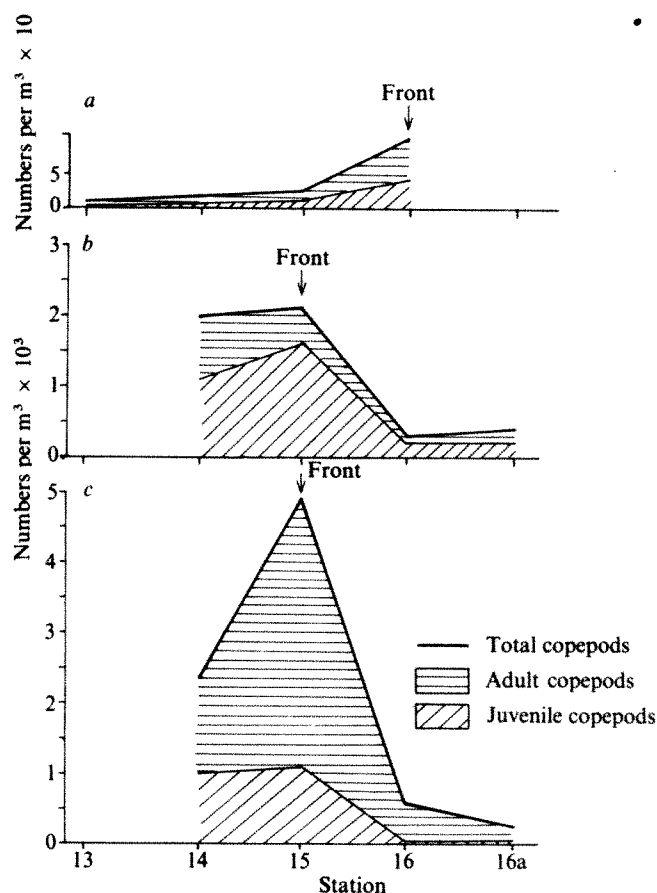


Fig. 5 Distribution of adult and juvenile copepods along the reference line indicated in Fig. 1 during *a*, March 1978; *b*, May 1978; and *c*, July 1978.

waters is not derived from discharges from the land but seems to be produced *in situ*, possibly by zooplankton (Turley, unpublished results)—in accordance with this, zooplankton were found to be concentrated at the front. Surveys in March, May and July 1978 found that total numbers of zooplankton at the front were between two and eight times higher than at stations on either side. Not only were there greater numbers of zooplankton but zooplankton development seems to have been more advanced at the front than elsewhere. When copepod distribution was examined, juveniles were found to reach a peak in this position in March and May whereas in July the peak was largely formed by adults (Fig. 5). This requires substantiation but we believe that breeding of copepods at the front is about 30 days in advance of the norm elsewhere in Liverpool Bay.

Thus, our results, which will be reported in more detail elsewhere, point to the activities of the various living components advected into a front being mutually accelerating. The greater numbers of bacteria inshore and in the winter suggest that they are mainly dependent on particulate and dissolved organic matter brought in by land run-off and these high populations seem from microscopical observations to be associated with particulate matter. Novitsky and Morita^{16,17} have related small cell size in a marine *Vibrio* to starvation conditions: both greater bacterial size and high ATP/cellular carbon ratios might therefore indicate that supply of bacterial nutrients is greater inshore and at the front than offshore. The increased concentrations of urea and greater urea breakdown associated with the front are in agreement with this. Any tendency for bacterial activity to increase in parallel with liberation of extracellular products by phytoplankton, as observed by Herbland¹⁸ in upwelling waters, is presumably masked by effects of terrigenous organic material, but the coincidence of peaks in both heterotrophic and phototrophic breakdown of urea in May (Fig. 4) suggests that intensification of bacterial activity occurs at

the front when there is a large phytoplankton population. It may also be significant that the mean bacterial size was greatest during the phytoplankton bloom. The high ATP/cellular carbon ratios at the time of transition from the bacteria to algal-dominated state, suggest that the microbial population was then in a particularly active state of catabolism¹⁹.

Zooplankton may be advected into the front and may also actively congregate there but then, because of the abundance of food, first in the form of particulate matter and bacteria and later in the form of phytoplankton, show accelerated development. Zooplankton grazing and excretion would in turn be expected to enhance the rate of recycling at the front.

We thank Drs C. P. Spencer, P. Foster and J. Simpson for advice on the chemistry and physical oceanography of fronts, Mr S. Czitrom for permission to quote his unpublished observations, Dr J. Beardall for the chlorophyll determinations, Mr T. Kendaris and Mr S. O. V. Scrope-Howe for assistance with the zooplankton work, and the officers and crew of R.V. Prince Madog for help. This work was in part supported by grant no. GR3/3091 from the NERC.

Received 25 September 1980; accepted 8 January 1981.

1. Pingree, R. D., Pugh, P. R., Holligan, P. M. & Forster, G. R. *Nature* **258**, 672-677 (1975).
2. Beardall, J. *et al. Biol. Contemp.* **5**(4), 163-175 (1978).
3. Okubo, A. in *Oceanic Fronts in Coastal Processes* (eds Bowman, M. J. & Esaias, W. E.) 23-28 (Springer, Berlin and New York, 1978).
4. Foster, P., Hunt, D. T. E., Pugh, K. B. & Foster, G. M. *J. exp. mar. Biol. Ecol.* **29**, 303-313 (1977); **34**, 55-71 (1978).
5. Daley, R. J. & Hobbie, J. E. *Limnol. Oceanogr.* **20**, 875-882 (1975).
6. Jones, J. G. & Simon, M. *J. appl. Bact.* **39**, 317-329 (1975).
7. Hobbie, J. E., Daley, R. J. & Jasper, S. *Appl. envir. Microbiol.* **33**, 1225-1228 (1977).
8. Holm-Hansen, O. & Booth, C. R. *Limnol. Oceanogr.* **11**, 510-519 (1966).
9. Yentsch, C. S. & Menzel, D. W. *Deep Sea Res.* **10**, 221-231 (1963).
10. Doetsch, R. N. & Cook, T. M. *Introduction to Bacteria and their Ecology* (University Park Press, Baltimore, 1973).
11. Luria, S. E. in *The Bacteria* Vol. 1 (eds Gunsalus, J. C. & Stanier, R. Y.) 1-34 (Academic, New York, 1960).
12. Strickland, J. P. H. in *Chemical Oceanography* Vol. 1 (eds Riley, J. P. & Skirrow, J.) 477-610 (Academic, London, 1965).
13. Simpson, J. H. & Pingree, R. D. in *Oceanic Fronts in Coastal Processes* (eds Bowman, M. J. & Esaias, W. E.) 29-42 (Springer, Berlin and New York, 1978).
14. Butler, E. J., Knox, S. & Liddicoat, M. I. *J. mar. biol. Ass. UK* **59**, 239-250 (1979).
15. McCarthy, J. J. *Limnol. Oceanogr.* **15**, 309-313 (1970).
16. Novitsky, J. A. & Morita, R. Y. *Appl. envir. Microbiol.* **32**, 617-622 (1976).
17. Novitsky, J. A. & Morita, R. Y. *Mar. Biol.* **48**, 289-295 (1978).
18. Herbland, A. in *Upwelling Ecosystems* (eds Boje, R. & Tomczak, M.) 155-166 (Springer, Berlin and New York, 1978).
19. Brezonik, P. L., Browne, F. X. & Fox, J. L. *Wat. Res.* **9**, 155-162 (1975).

Two distinct mechanisms of fibroblast adhesion

P. A. Harper* & R. L. Juliano†

* Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada M5S 1A1

† Department of Pharmacology, University of Texas Medical School at Houston, PO Box 20708, Houston, Texas 77025, USA

The adhesion of cells to the connective tissue matrix is commonly thought to be governed by fibronectin, a pericellular glycoprotein with binding sites for cell surfaces, collagen and glycosaminoglycans. Here we report evidence that Chinese hamster ovary (CHO) cells possess an alternative mechanism for adhesion which is independent of fibronectin. Cells of a variant CHO clone called AD^F11 are defective in their ability to adhere to fibronectin-coated substrata, but can adhere to a substratum coated with SAM (substrate-attached material), a pericellular material produced by fibroblasts. The adhesion of wild-type CHO cells to fibronectin-coated substrata and adhesion of AD^F11 cells to SAM-coated substrata are differentially sensitive to proteolytic treatment. This suggests that there are two distinct adhesion mechanisms for CHO cells, only one of which is dependent on fibronectin.

† To whom correspondence should be addressed.

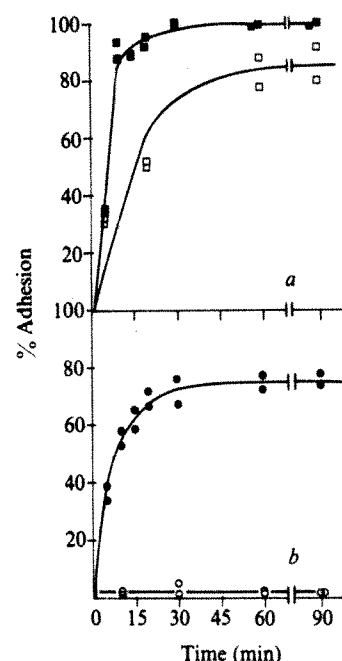
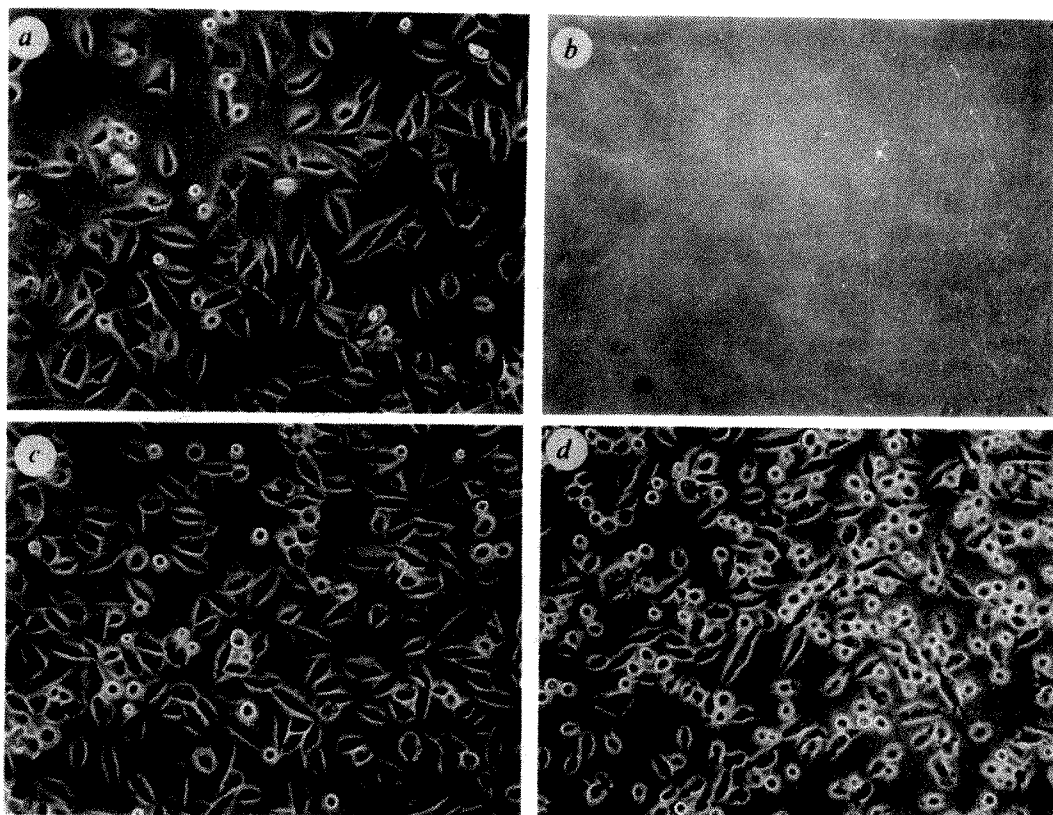


Fig. 1 Adhesion of WT and AD^F11 cells to SAM- and serum-coated substrata. The preparation of SAM- and serum-coated substrata, and the adhesion assay have been described elsewhere¹⁶. Aliquots of ³H-labelled suspension-adapted WT or AD^F11 cells were added to dishes coated with SAM or serum substrata and incubated at 37 °C and 5% CO₂. At various times, the dishes were removed from the incubator, washed extensively with PBS, pH 7.2, thus removing the non-attached cells and the fraction of radioactivity remaining on the plate due to the attached cells was determined. *a*: □, WT cells on serum; ■, WT cells on SAM. *b*: ○, AD^F11 cells on serum; ●, AD^F11 cells on SAM. Each point represents the mean of three determinations. The SAM used in these experiments was prepared from confluent human diploid fibroblasts as described elsewhere^{16,19}.

Variations in the capacity of a cell to make and break adhesive connections with other cells and with the extracellular matrix have been implicated in a number of important biological processes including organogenesis, wound healing and the invasive and metastatic behaviour of tumour cells¹. Recent studies of cell adhesion have focused on the role of the fibronectins, which are high molecular-weight glycoproteins found on cell surfaces, in the interstitial matrix and in plasma^{2,3}. Fibronectins can interact with the cell surface and with several types of macromolecules including collagen⁴⁻⁹, glycosaminoglycans^{10,11}, fibrin¹² and possibly glycolipids¹³; the binding properties of fibronectin, together with its pericellular location, suggest that this molecule provides an adhesive link between the cell surface and the extracellular matrix. Both the plasma and cellular forms of fibronectin can promote the adhesion and spreading of a number of transformed fibroblast cell lines on collagen-coated surfaces¹⁴, while cellular fibronectin can also restore a more 'normal' morphology to transformed fibroblasts¹⁵. It seems evident therefore, that the fibronectin-collagen complex is important in fibroblast adhesion. However, the possibility that macromolecular or cellular systems other than fibronectin may also be involved in adhesion has not been extensively explored. Here we present evidence for an alternative mechanism of cell adhesion independent of fibronectin.

As a tool for studying mechanisms of cell adhesion, we recently isolated and characterized a series of CHO cell variants (AD^F (adhesion variant) cells) which are markedly defective in their ability to adhere to fibronectin-coated collagen¹⁶; these cells also fail to adhere to serum-coated surfaces, as fibronectin is the major adhesion-promoting molecule in serum. Although AD^F cells have lost the ability to adhere to a fibronectin-coated substratum, these cells can attach and spread normally on

Fig. 2 Micrographs of WT and AD^vF11 on SAM- and serum-coated substrata. Dishes pre-coated with SAM or serum were prepared as described in Fig. 1 legend. Aliquots of AD^vF11 or WT cells were added to the pre-coated dishes and incubated at 37 °C and 5% CO₂. After 90 min the dishes were washed with PBS and the remaining attached cells were fixed with 0.1% glutaraldehyde in PBS. The cells were observed by phase contrast photomicroscopy. *a*, WT cells on serum; *b*, AD^vF11 cells on serum; *c*, WT cells on SAM; *d*, AD^vF11 cells on SAM.



substrata coated with other ligands which are capable of forming multivalent attachments to cell surfaces (for example, concanavalin A, polylysine). In addition, we found that AD^v cells can attach and spread on substrata coated with SAM, a pericellular substratum-attached material which is derived from fibroblasts and which has been reported to enhance the adhesion of some types of cells¹⁷.

When fibroblasts are removed from a substratum by treatment with chelating agents such as EGTA, a pericellular material is left on the surface. This material, known as SAM or 'footpads' is a complex mixture of macromolecules including proteins (fibronectin and actin are two prominent components)¹⁸, glycosaminoglycans and phospholipids^{19,20}, and is organized into small vesicular structures surrounded by amorphous or fibrous deposits^{21,22}. AD^v cells can adhere and spread on SAM-coated substrata, presumably because the cells can bind to one or more of the macromolecules in the SAM complex. We demonstrate that the adhesion of wild-type (WT) CHO cells to fibronectin-coated substrata and the adhesion of AD^v CHO cells to SAM-coated substrata are mediated by different mechanisms, which suggests that there are two separate systems for cell adhesion.

The adhesion behaviour of WT and AD^v cells on SAM-coated substrata and on serum-coated substrata is shown in Fig. 1. WT cells adhered rapidly and extensively (80–95%) to both types of substrata; by contrast clone AD^vF11 cells adhered readily (80%) to SAM-coated substrata, but failed to adhere (<3%) to serum-coated substrata. Figure 2 shows that the WT cells which attached to SAM- or serum-coated substrata went on to spread and assumed similar morphologies on the two surfaces. AD^vF11 also spread and attained a fairly normal CHO cell morphology on SAM substrata.

Treatment of SAM with several perturbing agents (Table 1*a*) did not impair its suitability as an adhesive substratum for WT and AD^vF11 cells; SAM treated with the non-ionic detergent Triton X-100 (TX-100), with chondroitinase ABC or with 4 M urea sustained cell adhesion as well as control SAM substrata. In these experiments >85% of sulphated glycosaminoglycans

Table 1 Perturbation of adhesion to SAM

<i>a. Treatment of SAM</i>		TX-100 (0.1%)	Ch ABC (0.5 U ml ⁻¹)	Thrombin (10 U ml ⁻¹)
Clone	Control			
WT	100	95.2 ± 5.9	98.6 ± 4.6	99.7 ± 3.0
AD ^v F11	100	94.2 ± 16.4	98.3 ± 1.5	97.5 ± 1.5
		Urea (4M)	Trypsin (5 µg ml ⁻¹)	
Clone	Control			
WT	100	99.8 ± 9.8	42.2 ± 8.6	
AD ^v F11	100	98.9 ± 10.0	12.5 ± 4.5	
<i>b. Treatment of cells</i>		Thrombin (10 U ml ⁻¹)	Trypsin (5 µg ml ⁻¹)	
Clone	Control			
WT	100	91.3 ± 14.1	100.6 ± 4.6	
AD ^v F11	100	65.7 ± 6.8	12.4 ± 3.3	

Effects of various agents on the adhesion of AD^vF11 and WT cells to SAM substrata. Substrates pre-coated with SAM were prepared as described in Fig. 1 except that, on removal of the detached fibroblasts, the dishes were further extracted with various agents. Thus dishes were incubated with 0.5 U ml⁻¹ chondroitinase ABC (Ch ABC, Miles) or 10 U ml⁻¹ thrombin (Sigma) in adhesion buffer, with 4 M urea or with 0.1% TX-100 in phosphate-buffered saline (PBS), pH 7.2, all for 40 min at 37 °C, or with 5 µg ml⁻¹ crystalline trypsin in PBS for 10 min at room temperature. After extraction, the dishes were washed with PBS and the attachment of WT and AD^vF11 after 90 min was determined as described in Fig. 1 legend. The data are expressed as the per cent of cells attached to treated plates against per cent attached to control SAM plates treated only with adhesion buffer. To evaluate the extent of the extraction procedures, SAM substrata derived from fibroblasts incubated with 10 µCi ml⁻¹ ³⁵S-sulphate to label GAGs, 10 µCi ml⁻¹ ³H-leucine to label proteins, or with 10 µCi ml⁻¹ ³H-oleic acid to label phospholipid, were similarly extracted and the fraction of radioactivity remaining on the plate determined (data in text). The data represent the mean and standard error of at least three experiments, each comprised of three or more determinations. Note that the SAM substrata probably also contain regions where the tissue culture dish is coated with serum, as the fibroblasts which produce SAM were originally grown in medium containing serum. Adhesion buffer comprised alpha MEM with 1 mg ml⁻¹ bovine albumin.

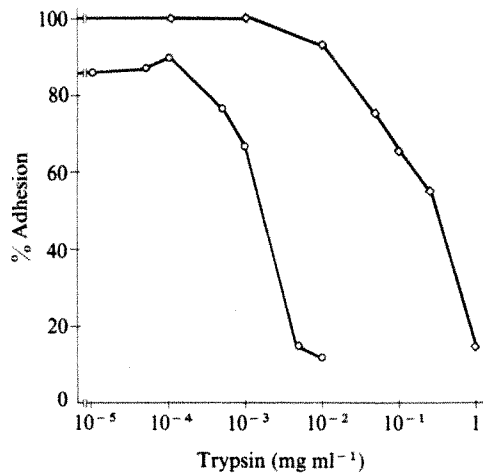


Fig. 3 The adhesion of trypsin-treated WT and ADⁱ cells to SAM- or serum-coated substrata. Suspension-adapted ³H-labelled WT or ADⁱF11 cells were treated with various concentrations of trypsin in PBS for 10 min at room temperature. Digestion was stopped by the addition of 50 volumes of ice-cold adhesion buffer containing 5 mg ml⁻¹ BSA followed by extensive washing of the cells; the cells were then assayed for their ability to attach to substrata coated with SAM or serum as described in Fig. 1 legend. To avoid the possibility of re-synthesis of surface components²³, the adhesion assay was limited to 30 min at 37 °C—sufficient time for the adhesion process to near completion. ◇, Trypsinized WT cells on serum substrata; ○, trypsinized ADⁱF11 cells on SAM substrata. Each point represents the mean of at least three determinations.

(GAGs) in SAM were removed by treatment with TX-100, chondroitinase ABC or 4 M urea; >75% of protein in SAM was removed by TX-100 or 4 M urea; >95% of the phospholipid was removed by TX-100; otherwise removal of GAG, protein and lipid was <25%. By contrast, treatment of SAM with low doses of trypsin (which removed ~75% of SAM protein) caused a marked reduction in its ability to support cell adhesion. In further experiments (Table 1b), treatment of ADⁱF11 cells with

thrombin or with low doses of trypsin markedly inhibited cell adhesion to SAM; by contrast WT adhesion to either SAM- or serum-coated substrata was unaffected by equivalent protease treatments of these cells.

These results suggest that ADⁱ cell attachment to SAM might be mediated by an interaction between cell-surface proteins and protein components in SAM. In addition, WT adhesion to serum-coated substrata and ADⁱ cell adhesion to SAM substrata might involve different cell components and mechanisms. This was confirmed by the experiments shown in Figs 3 and 4. Figure 3 shows that the cell-surface components involved in attachment of ADⁱ cells to SAM substrata were far more trypsin sensitive than the surface components involved in WT attachment to serum-coated substrata. The concentration of trypsin required to reduce adhesion of WT to serum substrata by 50% was 200 µg ml⁻¹, but only 2 µg ml⁻¹ for ADⁱ adhesion to SAM substrata.

The role of fibronectin in the adhesion of WT and ADⁱ cells was also evaluated using an anti-fibronectin antibody (Fig. 4). After treatment of the substratum with antibody, WT cells failed to adhere to fibronectin-coated collagen. However, antibody treatment of the SAM substratum did not impair the adhesion of WT or ADⁱF11 cells. Because the antibody effectively blocked fibronectin-mediated adhesion, its failure to block attachment to SAM implies that fibronectin is not directly involved in this process.

These results suggest the existence of two distinct mechanisms for adhesion in CHO cells, the first of which involves an association between fibronectin on the substratum and cell-surface components which are relatively trypsin resistant²³—this mechanism is utilized in WT CHO attachment to fibronectin-coated substrata. The second mechanism involves an association between trypsin-sensitive surface components and one or more substratum attached macromolecules other than fibronectin—this mechanism is used in ADⁱF11 attachment to SAM-coated substrata. It is likely that both mechanisms operate in WT cells, but that ADⁱ cells lack the fibronectin mediated mechanism. It seems premature at this stage to decide on the precise identity of the molecules involved in ADⁱ cell attachment to SAM. The results given in Table 1 and Fig. 4 suggest that neither GAGs nor fibronectin are involved; the possible involvement of fibroblast-secreted collagen or of other SAM components remains to be determined.

The authors thank Dr R. Rajaraman for the rabbit anti-human CIG. This work was supported by a grant from the NSF, and was carried out as part of a PhD thesis in the University of Toronto.

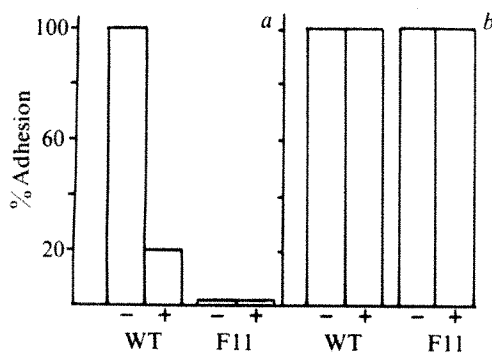


Fig. 4 Effect of anti-fibronectin antibody on adhesion. SAM-coated substrata were prepared as described in Fig. 1 legend. Collagen substrata (gelatinized rat-tail tendon collagen) were prepared as described elsewhere¹⁶ and were coated with 200 µg ml⁻¹ affinity-purified human serum fibronectin (CIG)²⁴. Both types of substrata were incubated in PBS with or without rabbit anti-human CIG diluted 1:10 for 90 min at room temperature. Thereafter the substrata were rinsed extensively with PBS and the adhesion of WT or ADⁱF11 cells during a 90-min period at 37 °C was measured, as described in Fig. 1 legend. Data represent the mean of three determinations. The data are normalized, based on the adhesion of WT cells to fibronectin-coated collagen (no antibody treatment) as 100%. *a*, Fibronectin-coated collagen substratum; *b*, SAM substratum. +, Pretreated with anti-fibronectin antibody; —, untreated.

Received 20 October; accepted 22 December 1980.

- Grinnell, F. *Int. Rev. Cytol.* **53**, 65–144 (1978).
- Vaheri, A. & Mosher, D. F. *Biochim. biophys. Acta* **51**, 1–25 (1978).
- Yamada, K. M. & Olden, K. *Nature* **275**, 179–184 (1978).
- Engvall, E., Ruoslahti, E. & Miller, E. J. *J. exp. Med.* **147**, 1584–1595 (1978).
- Kleinman, H. K. *et al. J. Biol. Chem.* **253**, 5642–5646 (1978).
- Hahn, L.-H. E. & Yamada, K. M. *Cell* **18**, 1043–1051 (1979).
- Ruoslahti, E., Hayman, E. G., Kuusela, P., Shively, J. E. & Engvall, E. *J. Biol. Chem.* **254**, 6054–6059 (1979).
- Balian, G., Click, E. M., Crouch, E., Davidson, J. M. & Bornstein, P. *J. Biol. Chem.* **254**, 1429–1482 (1979).
- Gold, L. I., Garcia-Pardo, A., Frangione, B., Franklin, E. C. & Pearlstein, E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4803–4807 (1979).
- Stathakis, N. E. & Mosesson, M. W. *J. clin. Invest.* **60**, 855–865 (1977).
- Perkins, M. E., Ji, T. H. & Hynes, R. O. *Cell* **16**, 941–952 (1979).
- Ruoslahti, E. & Vaheri, A. *J. exp. Med.* **141**, 497–501 (1975).
- Kleinman, H. K., Martin, G. R. & Fishman, P. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3367–3371 (1979).
- Yamada, K. M. & Kennedy, D. W. *J. Cell Biol.* **80**, 492–493 (1979).
- Yamada, K. M., Yamada, S. S. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1217–1221 (1976).
- Harper, P. A. & Juliano, R. L. *J. Cell Biol.* **87**, 755–763 (1980).
- Culp, L. A. *J. Cell Biol.* **63**, 71–83 (1974).
- Culp, L. A. in *Curr. Topics Membranes Transport* **11**, 327–396 (1978).
- Rollins, B. J. & Culp, L. A. *Biochemistry* **18**, 141–148 (1979).
- Cathcart, M. K. & Culp, L. A. *Biochemistry* **18**, 1167–1176 (1979).
- Revel, J. P., Hoch, P., Ho, P. & Ho, D. *Expl. Cell Res.* **84**, 207–218 (1974).
- Culp, L. A., Murray, B. A. & Rollins, B. J. *J. supramolec. Struct.* **11**, 391–400 (1979).
- Juliano, R. L. & Gagelang, E. *J. Cell Physiol.* **92**, 209–220 (1977).
- Engvall, E. & Ruoslahti, E. *Int. J. Cancer* **20**, 1–5 (1977).

Inter-hemispheric competition during postnatal development

Max Cynader*, Franco Leporé†
& Jean-Paul Guillemot‡

* Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J1

† Département du Psychologie, Université de Montréal, Montréal, Québec, Canada H3C 3J7

‡ Département du Kinanthropologie, Université du Québec à Montréal, Montréal, Québec, Canada H3C 3P8

Functional asymmetries between the two sides of the brain, a well documented phenomenon in species as different as frog and man^{1,2}, are thought to arise from genetically determined anatomical differences which, at least in humans, may be observed *in utero*³. Functional asymmetries can, however, be reversed after damage to one side of the brain⁴. Here we report that rearing of kittens with the optic chiasm sectioned and one eyelid sutured during postnatal development results in a functional asymmetry in the corpus callosum, a bidirectional pathway which inter-connects the visual cortices on the two sides of the brain. Visual input originating on the side of the brain ipsilateral to the sutured eye loses the ability to influence cells on the other side of the brain. Conversely, visual input originating on the side of the brain ipsilateral to the exposed eye markedly increases its influence in the other hemisphere.

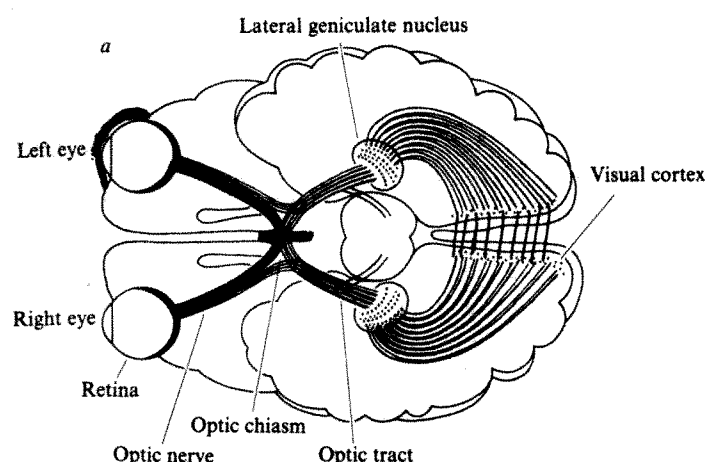
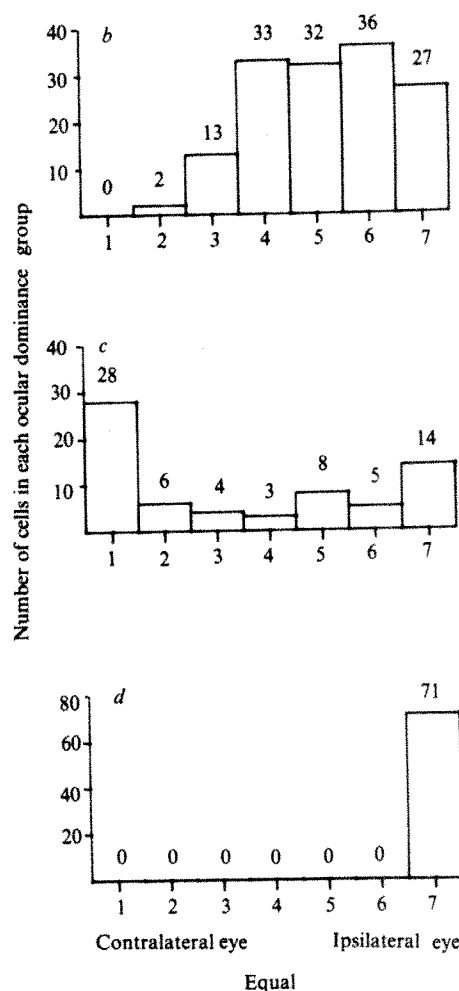


Fig. 1 *a*, General organization of the visual system in frontal-eyed mammals. Nerve fibres leave each eye and meet at the optic chiasm where fibres from the nasal hemi-retinae cross to the other side of the brain and those from the temporal hemi-retinae remain on the same side. Retinal inputs terminate in the LGN from where fibres project to the ipsilateral visual cortex. The corpus callosum inter-connects the visual cortices on the two sides of the brain. In the split-chiasm cats whose ocular dominance distribution is shown in *b*, the optic chiasm was subjected to midsagittal section when the animals were adults. This interrupts retinal input to the opposite side of the brain but leaves ipsilateral input intact. The abscissa of the ocular dominance distribution represents the relative excitatory contribution of the two eyes with increasing influence of the ipsilateral eye represented by successively higher numbers on a scale from 1 to 7 (ref. 16). The ordinate represents the number of cells encountered in each ocular dominance group. Most cells are more strongly influenced by stimulation of the ipsilateral eye in these cats, but many cells receive at least some input from the contralateral eye via the callosal pathway shown in *a*. *c*, 'Deprived hemisphere'. *d*, 'Exposed hemisphere': the ocular dominance distribution on the two sides of the brain in kittens in which the left eyelid was sutured at the same time as the chiasm was sectioned. In the hemisphere ipsilateral to the exposed eye all cells encountered were influenced solely by the exposed eye. In the other visual cortex, ipsilateral to the deprived eye, most cells received their major input from the eye on the opposite side of the brain despite the chiasm section.

The preparation used is shown in Fig. 1*a*. Four kittens were reared normally from birth until 21 days of age, when they were subjected to surgical section of the optic chiasm and at the same time the left eyelid was sutured^{5,6}. The kittens then received visual exposure in a normally lit environment until they were 4 months old. Optic chiasm section restricts visual input from one eye to one hemisphere by severing optic nerve fibres which cross to the opposite side of the brain. Suturing of the left eyelid prevents visual input from reaching the left visual cortex via the direct ipsilateral pathway through the lateral geniculate nucleus (LGN). The only route by which visual input can reach the left visual cortex in these kittens is a pathway from the right eye through the right LGN and visual cortex, then across the corpus callosum to the left visual cortex.

To compare the effectiveness of the direct visual route via the LGN with the indirect pathway described above in normally reared animals, we sectioned the optic chiasm in three adult cats and then recorded the visual responses of single cells at the border of cortical areas 17 and 18, where the corpus callosum is known to terminate⁷⁻⁹, using conventional methods described elsewhere¹⁰. The distribution of ocular dominance for cortical units found in these cats is shown in Fig. 1*b*, where the abscissa, numbered from one to seven, represents the relative excitatory input from each eye. After chiasm section in adult cats, most of the units encountered receive stronger input from the ipsilateral eye than the contralateral eye, indicating that the direct route



outlined above is more effective than the indirect pathway via the corpus callosum^{6,11}. Over 80% of the units encountered could however, be driven, to some extent, by stimulation of either eye. Despite the fact that the receptive field properties for the two eyes were determined by thalamic and callosal input respectively, visual response characteristics such as orientation, direction, and velocity selectivity were similar for a given neurone, regardless of which eye was stimulated. When one hemisphere was inactivated in these split-chiasm cats, all visual input to the remaining hemisphere was from the ipsilateral eye^{6,11}, which supports anatomical findings that the contralateral eye input reaches the cortex via the indirect route outlined above.

We recorded visual responses of single cells from the visual cortex on both sides of the brain in the four kittens in which one eyelid was sutured at the same time as chiasm section was performed. Multiple penetrations made oblique to the cortical surface ensured that a wide region of cortex near the 17/18 border was explored. The results are shown in Fig. 1d. On the side of the brain ipsilateral to the nonsutured eye, the 'exposed hemisphere', all units encountered were driven exclusively via stimulation of the exposed eye. No units were driven through the eye which had been sutured until the recording session began—this indicates a functional loss of input originating from the other side of the brain in these animals and mimics the situation in which that hemisphere is inactivated^{6,11}. Recording from the side of the brain ipsilateral to the previously deprived eye in these same kittens ('deprived hemisphere', Fig. 1c), we found that the effectiveness of input originating via the contralateral (exposed) eye was greatly enhanced relative to that found in adult split-chiasm cats. Inputs from this eye were clearly more effective than inputs from the deprived eye in driving most cortical cells. This indicates a substantial increase in the effectiveness of inputs originating on the side of the brain receiving visual stimulation during development relative to those from the deprived eye.

Our results show that the functional effectiveness of callosal connections can be markedly enhanced or diminished on different sides of the same brain. In effect, the corpus callosum has been transformed from a bidirectional pathway into a one-way route. Information originating in the 'exposed' hemisphere flows to the 'deprived' hemisphere but not in the reverse direction. These physiological findings are supported by the results of anatomical studies which show that the terminal field of the corpus callosum is markedly expanded in the 'deprived' hemisphere relative to the 'exposed' hemisphere¹². This induced functional asymmetry shows that the postnatal exposure history of the animal can modify the effectiveness of cortical output pathways in a similar manner to that by which it alters the effectiveness of inputs to the cortex¹³⁻¹⁵.

This research was supported by USPHS grant EY02248 and by grants MT5201 of MRC and A9939 and A9902 of NSERC of Canada.

Received 30 July; accepted 17 December 1980.

1. Dimond, S. J. & Blizard, D. A. *Ann. N.Y. Acad. Sci.* **299**, (1977).
2. Harnad, S., Doty, R. W., Goldstein, L., Jaynes, J. & Krauthamer, G. (eds) *Lateralization in the Nervous System* (Academic, New York, 1977).
3. Geschwind, N. *Fedn Proc.* **37**, 2263-2266 (1978).
4. Teuber, H. L. T. in *Physical Trauma as an Etiological Agent in Mental Retardation* (eds Angle, C. R. & Bering, E. Jr) 7-28 (NIH, Bethesda, 1970).
5. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26**, 1003-1017 (1963).
6. Berlucchi, G. *Brain Res.* **37**, 371-392 (1972).
7. Choudhury, B. P., Whitteridge, D. & Wilson, M. E. *Q. J. exp. Psychol.* **50**, 214-219 (1965).
8. Ebner, F. O. & Myers, R. E. *J. comp. Neurol.* **124**, 353-366 (1965).
9. Shatz, C. J. *comp. Neurol.* **171**, 229-245 (1977).
10. Cynader, M., Berman, N. & Hein, A. *Expl Brain Res.* **25**, 139-156 (1976).
11. Cynader, M., Dobbins, A., Gardner, J., Leporé, F. & Guillemot, J. P. *Soc. Neurosci. Abstr.* **5**, 781 (1979).
12. Cynader, M., Leporé, F., Guillemot, J. P. & Ferau, M. (in preparation).
13. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **28**, 1029-1040 (1965).
14. Cynader, M. & Mitchell, D. E. *Nature* **270**, 177-178 (1977).
15. Innocenti, G. & Frost, D. *Nature* **280**, 231-234 (1980).
16. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **168**, 106-154 (1962).

High-frequency transformation of the fission yeast *Schizosaccharomyces pombe*

David Beach & Paul Nurse

Biochemistry Laboratory and Genetics Group, University of Sussex, Brighton, Sussex BN1 9QG, UK

The fission yeast, *Schizosaccharomyces pombe*, has been used extensively for genetic studies but until now it has not been utilized as a host organism for DNA cloning. Here we describe a method for high-frequency transformation of a *leu*⁻ strain of this yeast with hybrid plasmids containing the *Saccharomyces cerevisiae* *LEU*²⁺ gene, a bacterial plasmid and either the *S. cerevisiae* 2 μ m plasmid or autonomously replicating sequences (*ars*)¹ derived from *S. pombe* DNA. Some of the plasmids contain unique restriction sites which make them suitable for the isolation of *S. pombe* genes, and they can also be used for the exchange of DNA between *S. pombe* and *S. cerevisiae*.

The plasmid, pJDB248, incorporates the *S. cerevisiae* *LEU*²⁺ gene which codes for β -isopropylmalate dehydrogenase, the bacterial plasmid pMB9 and the entire 2 μ m plasmid². It can be propagated in both *Escherichia coli* and *S. cerevisiae*, and transforms the *leu*⁻ *S. cerevisiae* strain MC16 (a leucine auxotroph lacking β -isopropylmalate dehydrogenase), at a frequency of 2×10^4 transformants per μ g DNA. We have tested the ability of this plasmid to transform *leu*⁻ auxotrophic strains of *S. pombe* to *LEU*⁺. Protoplasts were prepared from three strains, *leu* 1.32, *leu* 2.120 and *leu* 3.155. The procedure makes use of an extract of *Trichoderma harzianum*, Novo SP234 (Novo Enzymes), a rich source of 1-3 α glucanases (I. Isenberg, personal communication). Such *Trichoderma* extracts have been shown previously to digest the wall of *S. pombe*, to give protoplasts³. Protoplasts were washed three times in an osmotic buffer, 1.2 M sorbitol, to remove lytic enzymes, and incubated with plasmid DNA in the presence of CaCl₂ and polyethylene glycol essentially as described in ref. 2. After a further period of incubation in a leucine-supplemented medium, the protoplasts were plated directly onto minimal medium with 1.2 M sorbitol without using an agar overlay and incubated at 32°C for 4-5 days. The detailed method is given in Fig. 1 legend.

After this procedure only the *leu* 1.32 protoplasts give rise to colonies at a frequency greater than that of back-mutation. In the conditions used, pJDB248 transforms this strain to *LEU*⁺ at a frequency of 10^4 transformants per μ g DNA and 1 transformant per 10^4 viable regenerated protoplasts. The frequency of back-mutation of this strain is <1 in 10^7 . *S. cerevisiae* transformants of pJDB248 show frequent loss of the plasmid and consequent reversion to a *leu*⁻ phenotype. The stability of a *S. pombe* transformant of pJDB248 was tested and the proportion of *LEU*⁺ cells in a culture grown for 10 generations under selection in unsupplemented minimal medium was found to be 80%, indicating that the *S. pombe* transformants are also unstable. The presence of the plasmid in transformed *leu* 1.32 clones was demonstrated by preparing DNA from lysates of the yeast, and hybridization to a probe of the bacterial component of the plasmid using the Southern procedure⁴ (Fig. 1). Track 7 of the autoradiograph contains DNA from the untransformed *leu* 1.32 strain and shows no detectable hybridization. However, tracks 5 and 6 containing DNA from two separate pJDB248 transformants show hybridizing bands with electrophoretic mobilities corresponding to the supercoiled (rapidly migrating) and open circular or possibly dimeric (slowly migrating) forms of the plasmid.

These experiments show that *S. pombe* can be transformed at high frequency by the plasmid pJDB248, which contains no

S. pombe DNA. They also show that the *S. cerevisiae* *LEU* 2⁺ gene is expressed in *S. pombe* and that the *S. pombe* *LEU* 1⁺ gene must code for β -isopropylmalate dehydrogenase. In addition, the *S. cerevisiae* 2 μ m plasmid is probably capable of autonomous replication in *S. pombe*, as it is this part of pJDB248 which is likely to be responsible for plasmid replication. However, not all hybrid plasmids containing the 2 μ m plasmid and the *LEU* 2⁺ gene, and able to replicate in *S. cerevisiae*, can transform *S. pombe* at high frequency. An example is pJDB207 which contains only the 2.1-kilobase *Eco*RI fragment of the B form of the 2 μ m plasmid, and which transforms 2 μ m⁺ strains of *S. cerevisiae* at high frequency and has a high copy number (J. Beggs, personal communication). This plasmid transforms *S. pombe* *leu* 1.32 at low frequency (~20 transformants per μ g

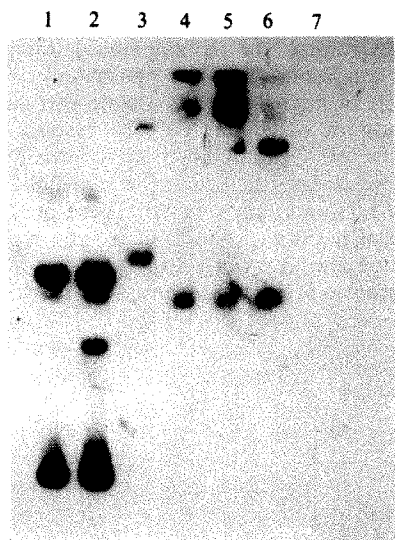


Fig. 1 Autoradiograph of ³²P-labelled pAC184 hybridized to a Southern transfer of undigested DNA of yeast transformants. Tracks 1 and 2, transformants of pDamP₄; tracks 3 and 4, transformants of pDamP₂; tracks 5 and 6, transformants of pJDB248; track 7, untransformed host strain. Cells to be transformed were grown to late exponential phase ($1-2 \times 10^7$ cells ml⁻¹) in leucine-supplemented (250 mg l⁻¹) minimal medium⁶. They were collected by centrifugation, washed with 20 mM citrate-phosphate, pH 5.6, and resuspended in the same buffer containing 40 mM EDTA, 0.14 M mercaptoethanol. After incubation at 32 °C for 30 min, the cells were resuspended in 20 mM citrate-phosphate, pH 5.6, containing 1.2 M sorbitol and 5 mg ml⁻¹ SP234 at 32 °C. After 45–60 min by which time protoplasts had formed, the cells were washed three times in 10 mM Tris-HCl, pH 7.6, 1.2 M sorbitol. Cells were resuspended at approximately 5×10^8 cells ml⁻¹ in 10 mM Tris-HCl, pH 7.6, 1.2 M sorbitol, 10 mM CaCl₂ and 10 μ g ml⁻¹ plasmid DNA. After incubation for 15 min at 25 °C, 10 vol of 10 mM Tris pH 7.6, 10 mM⁻ CaCl₂ 20% (w/v) polyethylene glycol 4000 (BDH) were added and incubated for a further 15 min. Finally the protoplasts were pelleted by centrifugation and resuspended at 5×10^7 cells ml⁻¹ in 10 mM Tris, pH 7.6, 10 mM CaCl₂, 5 μ g ml⁻¹ leucine, 0.5 mg ml⁻¹ yeast extract and 1.2 M sorbitol for 30 min at 32 °C. 0.2-ml aliquots (10^7 cells) were plated directly on to 2% agar minimal plates containing 1.2 M sorbitol without using an agar overlay and incubated at 32 °C. The percentage of protoplast regeneration, which was usually ~30%, was estimated using plates supplemented with leucine. Colonies grew up on the plates within 4–5 days. DNA was prepared from yeast by protoplast formation as previously described, followed by lysis in 10 mM Tris-HCl, pH 7.6, 10 mM EDTA, 2% SDS, 100 μ g ml⁻¹ proteinase K at 65 °C. After incubation for 30 min, the lysate was extracted with phenol and precipitated with 0.54 vol of isopropanol in 0.3 M sodium acetate. The precipitate was resuspended in 10 mM Tris-HCl, pH 7.6, 1 mM EDTA. The DNA samples were electrophoresed through a 1% agarose gel and transferred to a nitrocellulose filter⁴. The filter was hybridized with ³²P-pAC184 (500,000 d.p.m.) prepared by nick translation, washed and exposed to Kodak X-OMAT H film with an intensifying screen at -60 °C for several days.

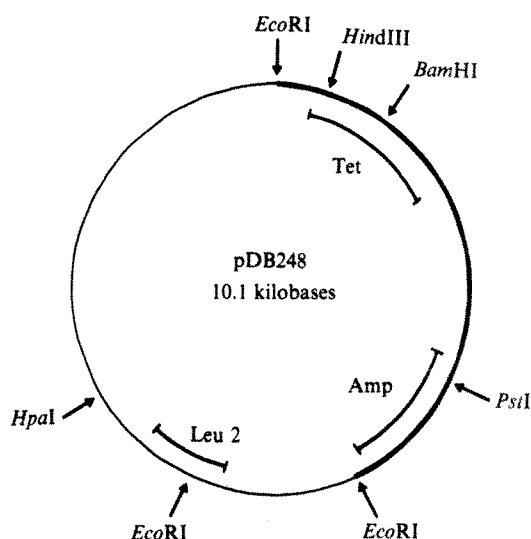


Fig. 2 Plasmid map of pDB248. This plasmid was derived from pJDB248 by partial digestion with *Eco*RI followed by ligation with a complete *Eco*RI digest of pBR322. *E. coli* (JA221) were transformed to *Leu* B⁺ and ampicillin resistance. pDB248 was one such transformant and contains pBR322 and the 4.4-kilobase *Eco*RI fragment of pJDB248. The *Eco*RI sites and unique *Pst*I, *Hind*III, *Bam*HI and *Hpa*I sites are shown. Amp; region coding for ampicillin resistance; Tet, region coding for tetracycline resistance.

DNA), possibly as a result of integration of the plasmid into the chromosome. Also pJDB219 (ref. 2), which contains the full-length 2 μ m plasmid and has a high copy number in *S. cerevisiae*, transforms *S. pombe* at an even lower frequency (~2 transformants per μ g DNA). The reasons for the low frequency of transformation are not apparent.

To have a plasmid capable of high-frequency transformation of *S. pombe* which is suitable for insertion of further genes (for example in the construction of an *S. pombe* gene bank), pJDB248 has been modified. The bacterial pMB9 sequence has been replaced with pBR322, and the 2.3-kilobase *Eco*RI fragment of the 2 μ m plasmid has been removed. The resulting 10.1-kilobase molecule, pDB248, carries both tetracycline and ampicillin resistance, and transforms *S. pombe* with the same frequency as pJDB248. Transformants of pJDB248 and pDB248 also have the same level of instability. pDB248 has unique restriction sites for *Pst*I, *Hind*III, *Bam*HI and *Hpa*I which makes it a suitable cloning vehicle (see Fig. 2).

To extend further the range of potential vectors which can be used to clone DNA in *S. pombe* we tested the transforming ability of eight previously constructed plasmids (pDam Y 1–8)⁵, containing the bacterial plasmid pBR325, the *LEU* 2⁺ gene and fragments of chromosomal DNA from *S. cerevisiae* which promote autonomous replication of plasmids in that yeast. These plasmids transformed *S. pombe* only at low frequency suggesting that these particular *S. cerevisiae* *ars* sequences are not recognized by *S. pombe*. For this reason new plasmids containing *S. pombe* rather than *S. cerevisiae* DNA were constructed by insertion of *Pst*I fragments of *S. pombe* DNA into pDam 1 (pBR325 containing the *LEU* 2⁺ gene). The plasmids were constructed by ligation of *Pst*I-digested pDam 1 (ref. 5), and a partial *Pst*I digest of *S. pombe* DNA. The ligated mixture was used to transform *E. coli* strain JA221 to chloramphenicol resistance. Forty ampicillin-sensitive colonies containing *Pst*I inserts into pDam 1 were picked, and mini-DNA preparations were made and each tested for ability to transform *S. pombe*. Of these 40 plasmids, 9 gave high-frequency transformation of the *leu* 1.32 strain (pDam P_{2–13}). However, the level of transformation obtainable with these plasmids is at least 10-fold lower than that with pJDB248. The transformants are unstable, with the proportion of *LEU*⁺ cells in cultures grown under selection ranging from 35 to 95%, depending on which pDamP

plasmid was present in the transformant. Free copies of the plasmid were also detected in the four transformants tested (Fig. 1). It is not clear whether the size difference of the plasmids obtained from the two independent pDam P₂ transformants (tracks 3 and 4) was the result of plasmid propagation in *E. coli* or *S. pombe*. We conclude that the nine *Pst*I fragments derived from *S. pombe* DNA contain autonomously replicating sequences as previously defined¹. It is uncertain whether these *ars* fragments contain true origins of replication from the *S. pombe* chromosome or are sequences with no chromosomal replicative function, but which happen to be able to 'support' the plasmid in the *S. pombe* cell.

We have described a highly efficient system for transformation of *S. pombe* similar to that developed for *S. cerevisiae*. The plasmid pDB248 contains unique restriction sites which makes it suitable for isolation of *S. pombe* genes. It can also be used to exchange DNA between the two quite different types of yeast. The finding that pJDB248 and pDB248, which contain no *S. pombe* DNA, can replicate in this yeast suggests wide application of the *S. cerevisiae* 2 μ m circle for DNA cloning in other eukaryotic cells.

We thank Melanie Piper and Rachel McIntosh for technical assistance, and Professor S. Shall for laboratory space. The plasmids pJDB248, pJDB207 and pJDB219 were a gift from Jean Beggs, and *leu* strains were supplied by Urs Leupold. We thank Novo for their gift of SP234. This work was supported by the CRC (D.B.) and SRC (P.N.).

Received 27 October 1980; accepted 21 January 1981.

1. Stinchcomb, D. T., Struhl, K. & Davis, R. W. *Nature* **282**, 39–43 (1979).
2. Beggs, J. D. *Nature* **275**, 104–109 (1978).
3. Kopecká, M. *Folia Microbiol.* **20**, 273–276 (1975).
4. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
5. Beach, D. H., Piper, M. & Shall, S. *Nature* **284**, 185–187 (1980).
6. Nurse, P. *Nature* **256**, 547–551 (1975).

Oxygen-dependence of chromosomal aberrations in Fanconi's anaemia

Hans Joenje*, Fré Arwert*, Aldur W. Eriksson*,
Han de Koning† & Anneke B. Oostra*

* Institute of Human Genetics, Free University,
1007 MC Amsterdam, The Netherlands

† Department of Pediatrics, University Hospital, Leiden,
The Netherlands

Fanconi's anaemia (FA) is an autosomal recessive disorder characterized by a high frequency of 'spontaneous' chromosomal aberrations¹ and an increased risk of cancer². If, as seems plausible, the microscopically visible chromosomal aberrations in this disorder result from DNA or chromatin damage that would normally be repairable, the questions arise as to which step (or steps) in the repair process is deficient and whether the deficiency is intrinsic or the result of secondary factors. We report here that the frequency of chromosomal aberrations in FA lymphocyte cultures is positively related to oxygen tension and suggest that the site primarily affected by the FA mutation is in the complex system of defence (that is, protection and repair) against the genetic toxicity of oxygen.

Whole-blood cultures stimulated by phytohaemagglutinin were incubated for 72 h at various oxygen tensions and screened for chromosomal aberrations. The range of oxygen tensions, 2–50% oxygen, was determined by depression of the mitotic index of both normal and FA cells in more extreme conditions. The frequency of chromosomal aberrations as a function of oxygen tension was found to be markedly different for normal subjects and FA patients.

For normal cells (Fig. 1) the frequency of aberrations was found to be almost constant at a mean of 2 (range 0–4) breaks per 100 metaphases. For FA cells, however, the frequency of aberrations increased dramatically with oxygen tension.

These observations show that FA cells are much more sensitive to cytogenetic oxygen toxicity than normal cells and indicate a role of oxygen in the process leading to chromosomal aberrations. Obvious candidates for oxygen-dependent chromosome damaging agents are the partially reduced ('activated') forms of oxygen, that is, superoxide-free radicals (O_2^-), hydrogen peroxide (H_2O_2), hydroxyl free radicals ($OH\cdot$) and singlet oxygen ($^1O_2$). These 'oxygen radicals' are formed as a result of normal oxygen metabolism³ and could contribute to chromosomal damage in two ways (Fig. 2): (1) by chemical interaction with the chromatin, either directly⁴ or by initiating lipoperoxidation, the products of which could interact with chromatin, resulting in damage⁵; (2) by interfering with DNA repair enzymes, directly⁶ or via lipoperoxidation products⁷.

In normal conditions the above pathways do not lead to alarming levels of microscopically visible chromosomal aberrations because of a cellular defence system consisting of antioxygenic factors (enzymes such as superoxide dismutases, catalase and peroxidases as well as low molecular-weight free-radical scavengers and antioxidants). These prevent excessive accumulation of harmful products of oxygen metabolism and lipoperoxidation and thus protect the genome; the last step of the defence process presumably consists of DNA or chromatin

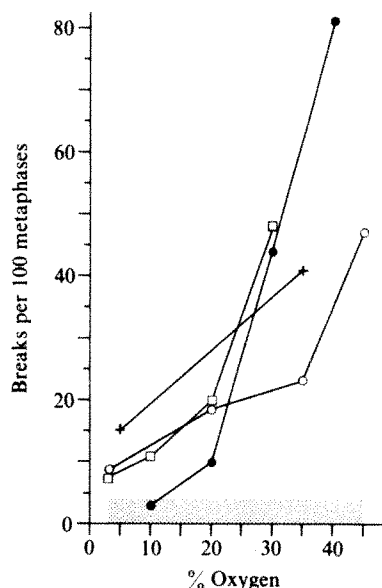


Fig. 1 Chromosomal aberrations as a function of oxygen tension. Whole blood cultures (0.5 ml heparinized venous blood plus 6 ml Ham F10 medium containing 15% fetal calf serum, phytohaemagglutinin and antibiotics, in glass infusion flasks with a total volume of 45 ml) were gassed with mixtures containing 1% CO_2 and the indicated percentages of oxygen balanced with nitrogen. After 72 h at 37 °C chromosome preparations were made according to an established procedure¹¹. 100 Giemsa-stained metaphases were viewed, on coded slides, for chromosomal aberrations. Metaphases with <40 chromosomes, tetraploids and endoreduplications were not scored. Chromatid and isochromatid breaks were scored as single breaks; dicentric, triradial and quadriradial as two breaks; more complex interchanges as the lowest possible number of breaks required for their reconstruction. Gaps were excluded. Control values, obtained from three individuals on three different occasions, lie within the shaded area. The symbols indicate different (unrelated) FA patients: ● (J.d.W.), a 20-year-old male; □ (P.W.), an 8-year-old male; ○ (R.H.), an 11-year-old male, and + (G.W.), a 10-year-old female.

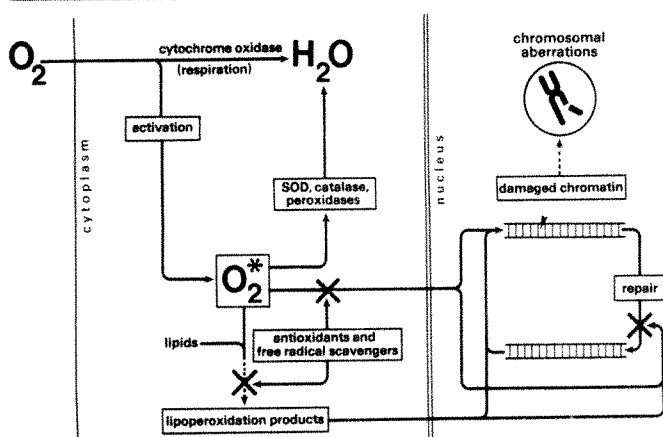


Fig. 2 Possible pathways in the genetic toxicology of oxygen. O_2^* , activated oxygen species.

repair reactions which remove damage due to products that have slipped through the system. In normal cells, the antioxygenic defence is still able to cope with mildly hyperoxic conditions, whereas in FA cells it seems to fail even in normal culture conditions.

Received 15 September 1980; accepted 14 January 1981.

- Schroeder, T. M., Anschutz, F. & Kopp, A. *Hum. Genet.* **1**, 194–196 (1964).
- Swift, M. *Nature* **230**, 370–373 (1971).
- Fridovich, I. *Science* **201**, 875–880 (1978).
- Cerutti, P. A. & Remsen, J. F. in *DNA Repair Processes* (eds Nichols, W. W. & Murphy, D. G.) 147–166 (Symposia Specialists, Miami, 1977).
- Freese, E. B., Gerson, J., Taber, H., Rhaese, H. J. & Freese, E. *Mutat. Res.* **4**, 517–531 (1973).

- Halliwell, B. *Cell Biol. int. Rep.* **2**, 113–128 (1978).
- Chio, K. S. & Tappel, A. L. *Biochemistry* **8**, 2827–2832 (1969).
- Poon, P. K., O'Brien, R. L. & Parker, J. W. *Nature* **250**, 223–225 (1974).
- Remsen, J. F. & Cerutti, P. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2419–2423 (1976).
- Fujiwara, Y., Tatsumi, M. & Sasaki, M. *J. molec. Biol.* **113**, 635–649 (1977).
- Moorhead, P. S., Nowell, P. C., Meilman, W. J., Battips, D. M. & Hungerford, D. A. *Exp. Cell Res.* **20**, 613–616 (1960).

Immunity to *Plasmodium chabaudi adami* in the B-cell-deficient mouse

James L. Grun & William P. Weidanz

Department of Microbiology and Immunology, Hahnemann Medical College and Hospital, Philadelphia, Pennsylvania 19102, USA

Immunity to malaria has a multicomponent basis which requires the participation of both T- and B-lymphocyte systems¹. Previous studies have suggested that the T-lymphocyte system has an essential role in 're-infection immunity' to malaria, but that B cells and/or their products are necessary for the host to survive acute infection and to clear the blood of parasites during chronic malaria^{2–4}. Thus, B-cell-deficient mice and chickens died of fulminant malaria when infected with *Plasmodium yoelii* and *Plasmodium gallinaceum*, respectively^{2,3,5}, but when their acute infections were controlled with subcurative chemotherapy, B-cell-deficient hosts developed chronic low-grade infections and resisted challenge with homologous parasites^{2,4}. In contrast, athymic nude mice failed to control their endogenous *P. yoelii* infection after the termination of drug therapy unless they had been thymus grafted before initiation of acute infection³. We now report that *Plasmodium chabaudi adami* (556KA) infection in B-cell-deficient mice results in an activation of a T-cell-dependent immune mechanism which terminates acute malaria in a similar way to that seen in immunologically intact mice. Furthermore, these immunized B-cell-deficient mice were resistant to homologous challenge infection as well as infections initiated with *Plasmodium vinckei*, but not with *P. yoelii* and *Plasmodium berghei*.

Age- and sex-matched hybrid mice (BALB/c × C57BL/10 and C57BL/6 × CBA) as well as *nu/nu* and *nu/+* mice on a BALB/c background were bred and reared in our animal facilities. Selected litters of mice were rendered B-cell deficient

by lifelong treatment with goat anti-mouse IgM serum (anti- μ) using methods described elsewhere⁴. Anti- μ treated mice, in comparison with immunologically intact mice, had no detectable levels of IgM and were deficient in spleen cells possessing membrane-bound immunoglobulin. B-cell-deficient mice continued to circulate goat anti- μ in their sera for prolonged periods of time and cells from these animals failed to produce specific antibody after immunization with sheep erythrocytes, or dinitrophenol coupled to keyhole limpet haemocyanin. Spleenocytes from treated animals did not respond to the B-cell mitogen lipopolysaccharide, but showed normal responsiveness to the T-cell mitogen, concanavalin A. These mice were as readily sensitized to strong contact allergens, for example, oxazolone and dinitrofluorobenzene, as normal control mice. Prolonged treatment with anti- μ activated macrophages locally within the peritoneal cavity, as shown by enhanced spreading and increased phagocytic activity when these cells were tested *in vitro*. However, there was no systemic activation of macrophages after this treatment because colloidal carbon was cleared from

Table 1 Selected parameters of immunological reactivity in B-cell-deficient mice compared with immunologically intact control mice, 45 d after infection with *P. chabaudi*

Treatment*	No. of animals	Serum IgM [†] level	Circulating [†] goat anti-IgM	Mean IFA titre [‡] <i>P. chabaudi</i>
Anti- μ	8	< 50 $\mu\text{g ml}^{-1}$	+	< 4
—	7	> 50 $\mu\text{g ml}^{-1}$	—	256

8-week old mice were infected i.v. with 10^5 parasitized erythrocytes. * (C57BL/6 × CBA)F₁ mice were injected i.p. three times weekly from birth with goat anti-IgM.

[†] Determined by Ouchterlony analysis⁴.

[‡] Titres were determined on serial fourfold dilutions of serum developed with fluorescein-labelled anti-mouse immunoglobulin and expressed as the reciprocal of the highest dilution of serum showing fluorescence⁴.

the blood of treated mice at the same rate as observed in untreated mice (data not shown).

Experimental infections with *P. c. adami* were initiated using a stablate of material originally obtained from D. Wyler (NIH). Mice were infected either intraperitoneally (i.p.) with 10^6 parasitized erythrocytes or intravenously (i.v.) with 10^5 parasitized erythrocytes. Infections were followed by microscopic examination of thin blood films prepared from tail blood and stained with Giemsa stain.

Infections with 10^5 *P. c. adami* produced a bimodal curve of non-fatal parasitaemia in both B-cell-deficient and immunologically intact mice (Fig. 1a). Following the second peak of parasitaemia, B-cell-deficient mice developed chronic low-grade infections as shown by positive blood films. Sub-inoculation of heparinized blood from chronically infected B-cell-deficient mice 50 days after infection to previously uninfected mice produced typical *P. c. adami* infections, indicating that the virulence of the parasites had not been attenuated by passage in the B-cell-deficient host. So far, acute infections with *P. c. adami* in more than 90 B-cell-deficient mice of assorted genotypes have been resolved successfully with no fatalities.

As treatment with anti- μ does not result invariably in immunosuppression^{6,7}, animals in this experiment were evaluated for their immunological competence 45 days after initiation of infection. As shown in Table 1, the sera of B-cell-deficient mice lacked detectable levels of IgM and serum antibodies reactive with *P. c. adami*. In addition, anti-mouse IgM was detected in the sera of all B-cell-deficient mice, demonstrating

Table 2 Specificity of immunity induced by *P. chabaudi* infection in B-cell-deficient mice

Group	No. of animals	Challenge infection	No. resistant No. challenged
A	5	<i>P. chabaudi</i>	5/5*
B	6	<i>P. vinckei</i>	6/6*
C	8	<i>P. berghei</i>	0/8†
D	5	<i>P. yoelii</i>	0/5†

6–8-week old (BALB/c × C57BL/10)F₁ mice were originally infected i.v. with 10^5 *P. chabaudi*-parasitized erythrocytes.

Challenge infections were initiated i.v. with 10^6 parasites 45 d after infection with *P. chabaudi*.

* Parasitaemias in challenged animals did not exceed 0.5%.

† All mice developed fulminant malaria and died within 3 weeks of challenge. Infections in previously uninfected B-cell-deficient mice followed a similar course.

that these animals remained suppressed although they had received intensive antigenic stimulation.

The ability of B-cell-deficient mice to control endogenous parasites was one indicator of immunity, but demonstration of specificity required challenge with exogenous parasites. Thus, B-cell-deficient mice which had previously terminated acute *P. c. adami* infections were challenged with assorted plasmodial species. The data (Table 2) demonstrate that antibody-independent immunity resulting from *P. c. adami* infection was of limited specificity, that is, immunized mice were protected against challenge infection with either *P. c. adami* or the virulent *P. vinckei*. However, *P. berghei* and the normally avirulent 17X strain of *P. yoelii* produced fatal infections. The limited specificity of the immunity induced by *P. c. adami* is similar to that found by Cox⁸, who used immunologically intact mice in cross-protection studies. This finding remains to be explained, but it is possible that more than one immune mechanism must be activated to enable the host to resist infection with certain plasmodial species^{9–11}.

Because antibody-independent mechanisms of immunity were shown to be T-cell dependent in the model systems studied previously^{2,4}, athymic nude mice were infected with *P. c. adami* to determine if immunity to this parasite would develop in the absence of functional T cells. As shown in Fig. 1b, nude mice were unable to terminate *P. c. adami* infection and had significant parasitaemias resulting in death 7 to 9 weeks post-infection. Thus, the development of immunity to *P. c. adami* infection is dependent on the function of an intact thymus-dependent lymphocyte system.

The most interesting aspect of this study is that acute infection with *P. c. adami* was controlled by antibody-independent immune mechanisms which seem to be T-cell dependent. This finding contrasts with previous results^{2,3,5} in which survival of the host following infection with *P. yoelii* and *P. gallinaceum* was dependent on B cells (presumably antibodies). Although the nature of the antibody-independent mechanisms responsible for immunity to acute *P. c. adami* infections in B-cell-deficient mice has not been determined, it is possible that activated macrophages play a significant part. Macrophage activation has been shown to occur during murine malaria¹² and it has been suggested that factors secreted by activated macrophages may stimulate both T cells and natural killer cells to produce lymphokines and inhibitory factors, respectively¹⁰. In addition, activated macrophages may produce inhibiting factors which directly kill certain intracellular protozoa¹³. Consistent with previous observations² concerning chronic plasmodial infections in B-cell-deficient hosts was the observation that B-cell-deficient mice were unable to clear *P. c. adami* from their blood during chronic infection despite their resistance to exogenous challenge. This supports the concept of Allison and Clark⁹ that immunity to malaria is mediated by multiple mechanisms of resistance. Particular mechanisms may predominate in infections caused by certain haemoproteoan parasites and/or

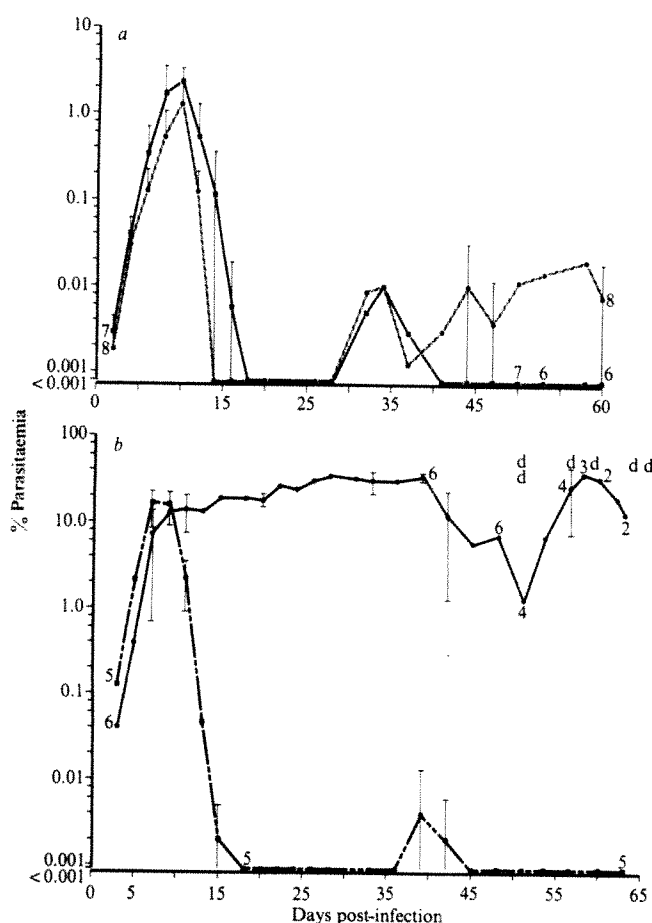


Fig. 1 a, The course of *P. chabaudi* infections in eight B-cell-deficient (●) and in seven immunologically intact mice (■). 9-week old (C57BL/6 × CBA)F₁ mice were infected i.v. with 1×10^5 parasitized erythrocytes. b, The course of *P. chabaudi* infections in six athymic nude mice (●) and in five euthymic littermates (■). Mean parasitaemias are shown for 7-week old mice infected i.p. with 1×10^6 parasitized erythrocytes.

may have special significance at a particular time during the course of disease. The elucidation of these mechanisms could provide targets for development of a potential 'malaria vaccine'.

We acknowledge the financial support of the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and the US PHS research grant AI-12710-4.

Received 6 October; accepted 17 December 1980.

1. Cohen, S. *Proc. R. Soc. B* **203**, 223-345 (1979).
2. Rank, R. G. & Weidanz, W. P. *Proc. Soc. exp. Biol. Med.* **151**, 257-259 (1976).
3. Roberts, D. W., Rank, R. G., Weidanz, W. P. & Finerty, J. F. *Infect. Immunity* **16**, 821-826 (1977).
4. Roberts, D. W. & Weidanz, W. P. *Am. J. trop. Med. Hyg.* **28**, 1-3 (1979).
5. Weinbaum, F. L., Evans, C. B. & Tigelaar, R. E. *J. Immun.* **117**, 1199-2005 (1976).
6. Manning, D. D. *J. reticuloendothelial Soc.* **18**, 63-86 (1975).
7. Dwyer, J. M., Rosenbaum, J. T. & Lewis, S. J. *exp. Med.* **143**, 781-790 (1976).
8. Cox, F. E. G. *Bull. Wild Hlth Org.* **53**, 325-336 (1970).
9. Allison, A. C. & Clark, I. A. *Am. J. trop. Med. Hyg.* **26**, 216-221 (1977).
10. Allison, A. C., Christensen, J., Clark, I. A., Elford, B. C. & Eugui, E. M. in *The Spleen and its Role in Parasitic Infections* (ed. Torrigiani, G.) 151-177 (Karger, Basel, 1978).
11. Eugui, E. M. & Allison, A. C. *Bull. Wild Hlth Org.* **57**, Suppl. 1, 231-238 (1978).
12. Shear, H. L., Nussenzweig, R. S. & Bianco, C. J. *exp. Med.* **149**, 1288-1298 (1979).
13. Clark, I. A. *Parasite Immun.* **1**, 179-196 (1979).

Immunotoxins: hybrid molecules of monoclonal antibodies and a toxin subunit specifically kill tumour cells

Hildur E. Blythman, Pierre Casellas, Olivier Gros, Pierre Gros, Franz K. Jansen, Francis Paolucci, Bernard Pau & Hubert Vidal

Centre de Recherches Clin Midy, Rue du Professeur Joseph Blayac, 34082 Montpellier Cedex, France

Several attempts to attack tumours in experimental systems have been made using conjugates of chemotherapeutic agents or potent toxins with antibodies (immunotoxins)¹⁻⁸. *In vitro* studies have been highly successful, showing target specificity of a high order in some cases⁶⁻⁸. However, so far, such conjugates have been inadequate *in vivo*, probably for two main reasons. First, conventional heteroclonal antibodies are perhaps inappropriate, because purification by biochemical methods leaves a large amount of non-antibody γ -globulins. The use of monoclonal antibodies may overcome this problem. Second, when whole toxins have been conjugated to antibodies there has been a strong residual nonspecific cytotoxicity due to the binding capacity of a subunit, the B-piece of the toxin. (Diphtheria toxin or ricin consist of two polypeptide subunits. The A-piece is responsible for inhibition of protein synthesis on ribosomes, and the B-piece binds to galactose residues on the cell membrane and facilitates the transmembrane passage of the A-piece.) In the present work the problem of nonspecific binding by the B-piece has been circumvented by using the A-piece only as the toxin component of immunotoxins; these immunotoxins are active both *in vitro* and *in vivo*.

A monoclonal IgM antibody directed against the Thy 1,2 differentiation antigen of mouse T cells, also present on mouse leukaemia cells (WEHI-7), was used. After precipitation of the monoclonal IgM antibody with ammonium sulphate, the antibody was purified with successive filtrations on Sepharose 6B. It was then modified with 3-(2-pyridyldithio)propionic acid in the presence of water-soluble carbodiimide and coupled to highly purified A-chain of ricin⁸. The conjugate was purified on a Sephadex G-200 column from unconjugated A-chain. The resulting immunotoxin contained on average eight A-chains per molecule, about three of which were active in an acellular protein synthesis system. The IgM component of immunotoxin retained about 70% of its activity in a complement-dependent cytotoxicity test. Inhibition of protein synthesis (PI) of viable cells *in vitro* was used to evaluate the biological activity of the conjugates⁸.

After incubation of 2×10^5 WEHI-7 cells for 16 h with different immunotoxin concentrations, a 50% inhibition of ^{14} C-leucine uptake (PI_{50}) was obtained with a 10^{-10} M concentration (Fig. 1a). Free A-chain was much less active, needing a concentration 5,000 times higher ($PI_{50} = 5 \times 10^{-7}$ M) while ricin itself was about 25 times more active than immunotoxin ($PI_{50} = 2 \times 10^{-11}$ M). A nonspecific immunotoxin directed against dextran, an unrelated antigen, behaved as free A-chain. The specificity of immunotoxin calculated as the increase of concentration needed to kill the same cells by the A-chain alone corresponded to a factor of between 2,000 and 6,000 in different batches tested (Fig. 1b). The antibody alone without complement showed no significant effect (Fig. 1c). Unconjugated antibody mixed with free A-chain behaved as free A-chain (Fig. 1c). No effect was found with an anti-Thy 1,2 immunotoxin on Thy 1,2-negative cells (Fig. 1d), although these cells had a similar sensitivity to ricin and A-chain. Complete inhibition of immunotoxin activity

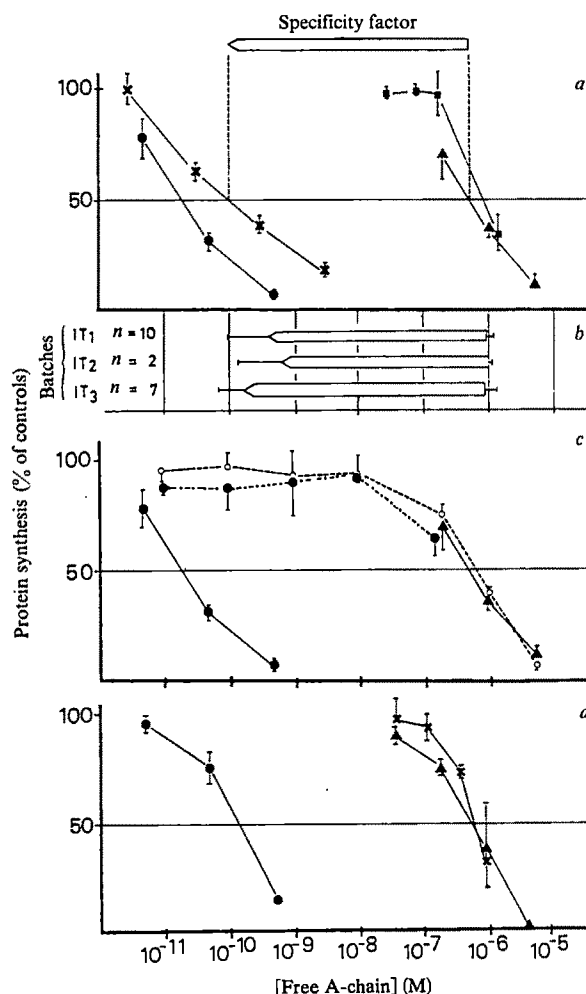


Fig. 1 Specific cytotoxic activity of anti-Thy 1,2 immunotoxin. *a*, 2×10^5 WEHI-7 cells (a mouse T leukaemia expressing Thy 1,2 antigen) were incubated for 16 h with different concentrations of the test substances and then pulsed with ^{14} C-leucine for 2 h (for more details see ref. 8). ●, Ricin; ▲, A-chain; ×, anti-Thy 1,2 immunotoxin; ■, anti-dextran immunotoxin. *b*, Specificity factors demonstrate the specific cytotoxicity of immunotoxins as compared with their toxic subunit, the A-chain alone, and were drawn as arrows with the tail at the concentration (M) of free A-chain which induces 50% protein synthesis inhibition and with the head at the concentration for specific protein synthesis inhibition by immunotoxin. The specificity factors of three batches of anti-Thy 1,2 immunotoxin are compared. *c*, Thy 1,2-positive cells (WEHI-7) were used to establish specificity controls for IgM alone without complement (---●---), IgM mixed with free A-chain (---○---), A-chain (---▲---) and ricin (---●---). *d*, Thy 1,2-negative cells (BC-3A) served as controls for the specificity of the Thy 1,2 immunotoxin (---×---) as well as sensitivity to ricin (---●---) and A-chain (---▲---).

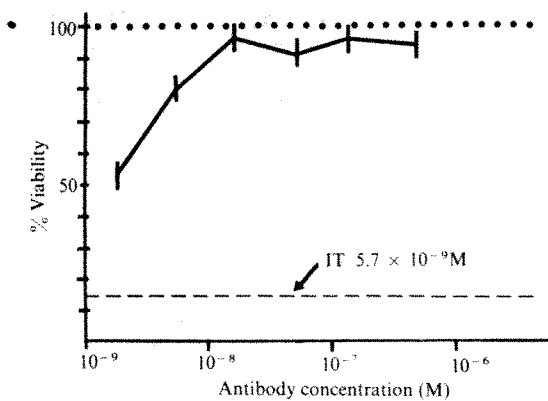


Fig. 2 Inhibition of immunotoxin activity by competition with unconjugated antibody. Increasing amounts of unconjugated antibody were added to a constant concentration of immunotoxin (5.7×10^{-9} M) which in standard tests gave 15% viability (dashed line). A threefold increase of unconjugated antibody (1.8×10^{-8} M) was needed to obtain 95% viability of target cells.

on Thy 1,2-positive cells (PI₈₅) could be obtained with a threefold molar excess of unconjugated antibody, suggesting that the binding capacity of immunotoxin is similar to that of unmodified antibodies (Fig. 2). Filter-sterilized conjugates stored in phosphate-buffered saline at -20°C for more than 6 months did not precipitate on thawing nor did they lose their immunotoxin activity.

In a first *in vivo* trial, WEHI-7 cells were injected intraperitoneally (i.p.) into BALB/c mice and each mouse was treated 1 h later with an i.p. immunotoxin injection of 30 μg conjugated A-chain, repeated 7 days later. This dose corresponded to about 1/10th of the LD₅₀ of free A-chain, which was 20 mg per kg (LD₅₀ of ricin was about 3,000 times lower, 7.5 μg per kg). With 6×10^5 cells, 9 of 10 control animals died within 39 days. Treatment with immunotoxin led to a significantly prolonged survival, as verified with the method of probits⁹. 50% survival was attained in the control group after 35 days and in the treated group after 44 days (Fig. 3b). However, with 2.4×10^5 cells no significantly prolonged survival could be obtained (Fig. 3a).

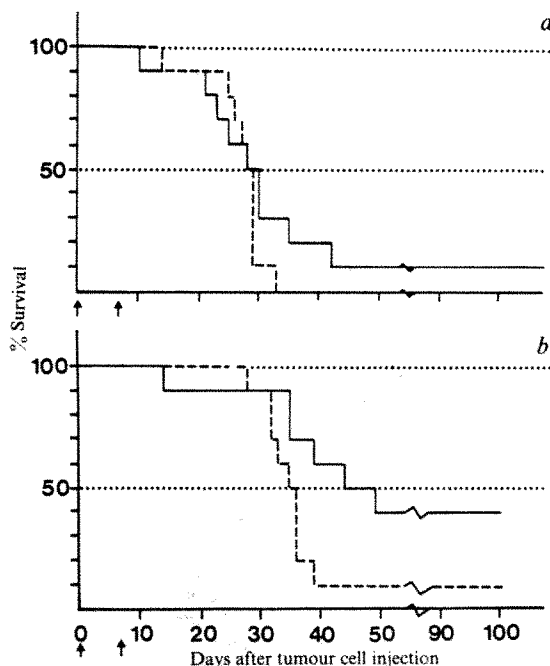


Fig. 3 *In vivo* test for immunotoxin activity. Groups of 10 BALB/c mice (Bomholtgaard) were injected i.p. with 2.4×10^5 WEHI-7 cells (a) or 6×10^5 WEHI-7 cells (b). After 1 h and 7 days, 30 μg of A-chain (bound to anti-Thy 1,2 antibody) were injected i.p. Survival was followed for 100 days.

The high specificity of immunotoxins consisting of monoclonal IgM antibodies (specificity factors of 2,000–6,000) may be attributed to the lack of non-antibody molecules or to the higher avidity of IgM over IgG antibodies. The nonspecific activity of immunotoxins on antigen-negative cells is very low and can be explained by the high purity of the A-piece replacing whole ricin. Such an effect was not convincingly demonstrated by Masuho *et al.* because their mouse cell system is insensitive to whole diphtheria toxin and cannot reveal contamination by the B-piece or whole toxin¹⁰. Specificity factors were not reported in the recent communications by Gilliland *et al.*¹¹ and Krollick *et al.*¹² because 50% nonspecific toxicity was not indicated.

Although a significant prolongation of survival time was achieved *in vivo* with our approach, the effect seemed less impressive when compared with the high potency of immunotoxin *in vitro*. This difference can perhaps be explained by the fact that the injected leukaemia cells represented less than 1% of all normal host cells expressing the Thy 1,2 antigen and thus the immunotoxin could have been absorbed by them. With repeated daily injections of immunotoxin compensating for its degradation or loss, better results might be expected. Although the immediate i.p. treatment of injected leukaemia cells with immunotoxin cannot be regarded as a model system of leukaemia treatment, it represents a highly sensitive test system. Whereas, *in vitro*, 50% survival of tumour cells is measured, *in vivo*, even single cell clones escaping the treatment will induce tumours some time later.

Immunotoxins retained most of their complement fixing activity but, as human complement is thought to be relatively inefficient *in vivo* for the destruction of tumour cells, the complement-independent activity of immunotoxins may represent a promising means of passive immunotherapy of cancer.

This work was supported in part by a grant from the Délégation Générale à la Recherche Scientifique et Technique, Paris, France. WEHI-7 and BC-3A cells were obtained from the Cell Distribution Center, the Salk Institute.

Received 10 October 1980; accepted 7 January 1981.

1. Mathé, G., Loc, T. B. & Bernard, J. C. *C. r. hebdomadaire Séances Acad. Sci., Paris* **246**, 1626–1628 (1958).
2. Ghose, T. & Blair, A. H. *J. natn. Cancer Inst.* **61**, 657–672 (1978).
3. Davies, D. A. L. & O'Neill, G. J. *Br. J. Cancer* **28**, Suppl. 1, 285–298 (1973).
4. Arnon, R. *Proc. Sero Symp.* **16**, 287–304 (1979).
5. Moolten, F. L., Zajdel, S. & Cooperband, S. R. *Ann. N. Y. Acad. Sci.* **277**, 690–699 (1976).
6. Thorpe, P. E. *et al. Nature* **271**, 752–755 (1978).
7. Youle, R. J. & Neville, D. M. Jr *Proc. natn. Acad. Sci. U.S.A.* **77**, 5483–5486 (1980).
8. Jansen, F. K. *et al. Immun. Lett.* **2**, 97–102 (1980); French Patent 78.27.838 (1978); U.S. Patent Application 79441 (1979).
9. Finney, D. J. in *Probit Analysis* 19–80 (Cambridge University Press, 1971).
10. Masuho, Y., Hara, T. & Noguchi, T. *Biochem. biophys. Res. Commun.* **90**, 320–326 (1979).
11. Gilliland, D. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4539–4543 (1980).
12. Krollick, K. A., Villemez, C., Isakson, P., Uhr, J. W. & Vitetta, E. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5419–5423 (1980).

Further molecular complexities of H-2 K- and D-region antigens

P. Démant & Dagmar Iványi

Division of Genetics, Antoni van Leeuwenhoek Huis, The Netherlands Cancer Institute, Amsterdam, The Netherlands

The K and D regions of the H-2 gene complex are highly polymorphic^{1,2} and control cell-surface structures involved in the H-2 restriction of cytotoxic T lymphocytes^{3,4}. Originally it was assumed that each of these regions controlled only one type of antigenic molecule, H-2K and H-2D, respectively^{2,5–7}. The finding that the D region controls H-2L^{8–10} and H-2M¹¹, as well as H-2D, molecules suggests that the system of H-2 antigens is more complicated. We demonstrate here previously unknown

H-2 molecules detected by anti-K- and D-region alloantisera in co-capping experiments. Three different types of molecule, distinct from Ia and Qa antigens, are distinguished in the products of the D^{dx} region, and two in the products of the D^k and K^d regions. Analysis of the antigenic heterogeneity of K- and D-region products may further understanding of the function of these regions in cell-mediated immunity.

Antibodies against the products of the K and D regions detect molecules of molecular weight 45,000 associated with a β_2 -microglobulin in chain¹²⁻¹⁴. This was demonstrated also for the H-2L¹⁵⁻¹⁷ and H-2M¹⁸ molecules. Alloantibodies against these antigens detect either specificities characteristic for a single allele, for example K^b or D^d (private specificities), or the common or cross-reactive specificities (public specificities) shared by products of several different alleles^{1,2,5,6,19}. Antisera against the public or private specificities were used to cap the respective K- and D-region products on T lymphocytes of C3H.LG and GRS ($H-2^{dx}$) and C3H.OH ($H-2^{oz}$) strains. After the capping procedure, the cells were tested with the antiserum used for capping as well as with other anti-private and anti-public antisera to see whether any molecules reactive with these antisera remained non-redistributed. Two techniques, indirect immunofluorescence and resistance to complement lysis (lyso-strip), were used as described earlier¹¹. The two assays never gave discordant results¹¹.

In earlier observations^{7,8,20}, capping with antisera against the public specificities removed all molecules reactive with antisera against the pertinent private specificity. We show that selected anti-public sera, however, fail to do so. Such sera were used to differentiate further the K- and D-region products. Table 1 summarizes the analysis of the H-2 molecules controlled by the $H-2^{dx}$ haplotype (K^d, I^f, S^f, D^{dx}) which was previously shown to control three molecules: H-2K^d, H-2D^{dx} and H-2L^{dx} (ref. 21). Data in Table 1 show that, using suitable antisera, five different H-2 molecules can be distinguished in this haplotype. Antiserum ASA-26B was made against the products of the D^{dx} region (the $H-2^{dx}$ and recipient have the same K^d region) and it contains two kinds of antibodies—one reactive with the H-2D^{dx} molecule and another with the H-2L^{dx} molecule²¹. The latter antibodies reacted with a public specificity common to H-2L^{dx} and some H-2.1-positive molecules produced by other haplotypes. Therefore, they could be removed from the ASA-26B by

absorption with C3H ($H-2^k$) cells leaving only anti-H-2D^{dx} antibodies. In capping experiments, the serum ASA-26B absorbed with (abs.) C3H (anti-H-2D^{dx}) did not cap all molecules detected by the antiserum ASA-26B (thus defining the H-2L^{dx} molecule, ref. 21 and Table 1). Antiserum ASA-26B abs. C3H did cap all molecules reactive with two other anti-public antisera, ASA-30C and ASA-14F, indicating that these sera contained no antibodies reactive with either H-2L^{dx} or K-region molecules. One of these sera, ASA-30C, however, failed to cap all molecules reactive with the ASA-26B abs. C3H and ASA-14F indicating that besides the H-2L^{dx} molecules (which are not capped by ASA-26B abs. C3H and ASA-14F) two other D^{dx} -region molecules can be distinguished: both can be capped by the ASA-26B abs. C3H and ASA-14F, but only one of them is capped by the ASA-30C. Similarly, the antisera D-31 and S3E which react with the K-region products of the $H-2^{dx}$ haplotype in capping experiments define two serologically different molecules. This is seen by the failure of the serum S3E to cap all molecules reactive with the serum D-31. Thus, two molecules can be distinguished in the K^d region: one which can be capped by the serum D-31 and S3E, and another which can be capped by D-31 but not by S3E. The anti-Ia antibodies are not responsible for the complexities of the K region described here as our assay does not detect Ia antigens on T lymphocytes^{8,11}. Moreover, these results were obtained with the K^d -region products in the cells carrying the $H-2^{dx}$ haplotype (K^d, I^f, S^f, D^{dx}) even after removal of anti-Ia antibodies by absorption (Table 1,*). The antiserum S3E cannot contain any anti-Ia antibodies, because both the donor and recipient are I^k .

Note that, although the genetic mapping of these molecules is only inferential as $H-2^{dx}$ is a natural recombinant, there is no doubt about the serological distinctness of these molecules and their H-2-like characteristics (ASA-26 and ASA-30 do not contain anti-Ia antibodies and in the strain GRS they do not react with Qa antigens²¹, and, as discussed above, the two K^d -region molecules are not Ia antigens).

A similar complexity was detected in the products of the $H-2^{oz}$ haplotype (K^d, I^d, S^d, D^k) (Table 2). The D^k -region products which are capped by anti- D^k -private serum ASA-11 or D-32 cannot be completely capped by the antiserum ASA-14. Similarly, the products of the K^d region which can be capped by the anti- K^d -private serum D-31 cannot be completely capped by

Table 1 Molecular relationships between H-2 specificities analysed by antibody-induced redistribution (capping) and resistance to lysis (lyso-strip) on T lymphocytes of the C3H.LG and GRS/Mtv-2- strains (both $H-2^{dx}$, K^d, I^f, S^f, D^{dx})

First antiserum	Second antiserum											
	ASA-26B		ASA-26B abs. C3H/Sn		ASA-30C		ASA-14F		D-31		S3E	
	IF	C	IF	C	IF	C	IF	C	IF	C	IF	C
None	100	100	100	100	100	100	100	100	100	100	100	100
ASA-26B	0	5	0	10	0	5	0	5	100	95	95	95
ASA-26B abs. C3H/Sn	100	90	0	5	0	5	10	10	100	100	100	100
ASA-30C	100	100	95	90	0	0	ND	75	ND	100	ND	100
ASA-14F	ND	100	ND	5	ND	5	ND	5	ND	100	ND	90
D-31	100	100	100	100	100	100	100	100	0	10	0	5
S3E	ND	95	ND	95	ND	75	ND	95	100	95*	0	5*
S3E + ASA-26B	0	0	ND	5	ND	5	ND	10	100	90*	0	5*

Strain C3H.LG was supplied by Dr M. Cherry. It carries the $H-2^{dx}$ haplotype which is indistinguishable serologically from that of GRS/A (P.D. and M. Cherry, unpublished data). The strain GRS/Mtv-2- was prepared from the GRS/A strain as described²⁸. It has the same H-2 haplotype as GRS/A and C3H.LG (data not shown). Results with the two strains were identical in cytotoxicity. Immunofluorescence was done predominantly on GRS/Mtv-2- cells. For the first antiserum, T lymphocytes were incubated with anti-H-2 serum + fluorescein isothiocyanate-labelled anti-IgG in capping conditions¹¹. The second antiserum was incubated in non-capping conditions¹¹ with addition of rhodamine-labelled anti-IgG or with complement. Antisera: ASA-26B (B10.D2 × A.CA)F₁ anti-GRS ($d \times f$) anti- dx (this antiserum ASA-26B contains anti-H-2.63, anti-H-2.1-like and non-H-2 antibodies²¹). These non-H-2 antibodies do not react with antigens of the C3H strain and its congenic lines (data not shown). Absorption with C3H or B10.A cells removes anti-H-2.1-like but not anti-H-2.63 activity²¹; ASA-30C (C3H × BALB/c- $H-2^{dm2}$)F₁ anti-BALB/cHe; ASA-14F ((A.SW × BALB/c)F₁ anti-A/Sn); D-31 (B10 × A)F₁ anti-B10.D2; and S3E (B10.RIII × C3H)F₁ anti-B10.AKM. Antiserum D-31 was supplied by the NIH serum bank. In some experiments NIH antiserum Ia.11,16 was used in place of D-31. This serum was prepared in the same strain combination and gave identical results. IF, immunofluorescence: % of cells with red label outside the green caps observed in capping experiments. C, cytotoxicity: redistribution results expressed as % of cells killed by the second antiserum and complement after incubation with the first antiserum + anti-IgG. Several serial dilutions were tested. ND, not determined.

* The same results were obtained with antiserum D-31 absorbed with B10.P cells (which absorb all activity with the strain B10.M, $H-2^f$, but not with C3H.OH, C3H.LG or GRS cells).

Table 2 Molecular relationship between H-2 private and public specificities measured by antibody-induced redistribution (capping) and resistance to lysis (lysostrip) on T lymphocytes of the C3H.OH ($K^d I^S D^k$) strain

First antiserum	ASA-11*		D-32		Second antiserum ASA-14		D-31*		S3E	
	D^k private		D^k public		K^d private		K^d public			
	IF	C	IF	C	IF	C	IF	C	IF	C
None	100	100	100	ND	100	90	100	100	100	100
ASA-11 (D^k private)	0	0	0	ND	ND	5	100	100	90	ND
D-31 (K^d private)	100	100	ND	ND	84	90	0	0	0	5
ASA-11 + D-31 ($D^k + K^d$ private)	0	5	0	ND	0	5	0	5	0	ND
D-32 + D-31 ($D^k + K^d$ private)	0	ND	0	ND	0	ND	0	ND	0	ND
ASA-14 (D^k public)	100†	100†	ND	ND	0	5	100	100	ND	ND
S3E (K^d public)	ND	100	ND	ND	ND	ND	ND	100	ND	5
ASA-14 + D-31 (D^k public + K^d private)	81†	100†	ND	ND	0	5	0	5	ND	ND
ASA-11 + S3E (D^k private + K^d public)	ND	5	ND	ND	ND	ND	90	100	0	5

For first and second antisera, see Table 1 legend. Antisera: ASA-11 (B10.AKM × C3H.B10)F₁ anti-C3H; D-32 (B10.A(2R) × C3H.SW)F₁ anti-C3H; ASA-14 (A.SW × BALB/c)F₁ anti-A/Sn; D-31 (B10 × A)F₁ anti-B10.D2; S3E (B10.RIII × C3H)F₁ anti-B10.AKM. The antisera D-31 and D-32 were supplied by the NIH serum bank. IF, immunofluorescence, and C, cytotoxicity; see Table 1 legend. ND, not determined.

* Several anti-public sera caused complete capping or patching of antigens detected by anti-private sera, for example ASA-1H capped for D-31b and D-5AF for ASA-11 (ref. 29).

† This reaction is not due to anti-Qa antibodies. ASA-11 was made by immunization against C3H/HeA cells. No Qa antigens were described in C3H strains. Moreover, we tested the tissue distribution of the antigens detected by this serum and found that liver and kidney tissue can absorb all activity of this serum with peripheral lymph node cells. Thus, the tissue distribution of H-2D1^k and H-2D2^k is like that of H-2 and not Qa.

the antiserum S3E. Thus, at least two different molecules can be distinguished among the D^k -region products. One can be capped by both ASA-14 and ASA-11 or D-32, the other only by the anti-H-2.32 sera (ASA-11, D-32) but not by ASA-14. Two distinct K^d products can be distinguished—one which can be capped by the anti-H-2.31 serum D31 as well as by the antiserum S3E, and another which can be capped only by the anti-H-2.31 serum and not by the antiserum S3E. No known Qa or TL antigens are involved in the complexities of the $D^k D^d$ -region antigens described here (see above and Table 2,†). The failure of some anti-public sera to cap all molecules reactive with the anti-private sera is not due to some intrinsic nonspecific properties of such anti-public sera. For example, the serum ASA-14 caps only one of the two D^k molecules (Table 2) but it can cap two different D^d products (Table 1).

Our data suggest, but do not prove, that the two D^k molecules described here may share the same private specificity, H-2.32. Alternatively, they could each carry a different unique specificity, but be indistinguishable because of the lack of suitable recombinants or mutants. The possibility that the antiserum ASA-14 does not react with an H-2.32 (ASA-11)-positive molecule, but with a molecule which co-caps with the H-2.32-positive molecules (as was suggested for some antigens on virally infected cells²²), also cannot be ruled out. The same interpretations can also be applied to the serological characteristics of the two different K^d -region molecules and the two types of D^d molecule capped by the ASA-26B abs.C3H. These uncertainties have to be resolved by immunoprecipitation studies. The relationship of the two D^k -region molecules to the H-2L molecules is unclear. Using different anti-H-2.32 and anti-H-2.1 sera we failed to confirm the original report²³ on the H-2.1-positive molecules detected after capping of H-2.32 molecules. Either the H-2.32-negative D^k -region products do not exist, or they carry their own 'unique' specificity and the antibodies against it contaminate the anti-H-2.32 sera.

Previously we demonstrated three different molecules in the D -region products of H-2^d (ref. 11). These are H-2L^d which cannot be capped by antisera against the specificity H-2.4, and H-2D^d and H-2M^d which can be capped by anti-H-2.4 sera, but the latter not by some anti-H-2.28 sera (S1E, (B10.RIII × C3H)F₁ anti-B10.A(2R)). Thus, together with the two K^d -region molecules, the H-2^d haplotype controls at least five serologically distinct H-2-like molecules.

It is not known whether the heterogeneity of the K - and D -region antigens described here reflects genetic complexity of these regions or is produced epigenetically. The former

explanation is supported by our observation that the mutant H-2^{dm1} does not express any detectable H-2M^d molecules¹¹, although even here post-translational mechanisms are not ruled out. Where several different H-2 molecules map into the same genetic region, the question of the respective genes becomes unclear. Until informative recombinants are found, these molecules would seem to be the products of pseudoallelic genes and the H-2 terminology which indicates, for example, that H-2D^d is 'allelic' or 'homologous' to H-2D^a, or H-2L^d is allelic or homologous to H-2L^{dx}, is based on apparent general serological similarities of the respective molecules, and may not be correct. Until suitable recombinants are found or other reliable genetic, biochemical or functional criteria of 'allelism' become available, it is very difficult to be certain about the relationships of molecules detected in the products of different regions. Therefore, we have used here a neutral terminology for the 'new' molecules, which makes no unwarranted assumptions about genetics and can be replaced by a more specific one when more information becomes available. Thus, the two molecules in the D^k region are named H-2D1^k and H-2D2^k, the two new D^d molecules are named H-2D1^{dx} and H-2D2^{dx}, and the two K^d -region molecules H-2K1^d and H-2K2^d.

It is also not certain whether each non-mutant haplotype produces all the known types of H-2 molecules. Furthermore, an H-2 molecule might be controlled by more than one gene⁹. A similar possibility was indicated for HLA (histocompatibility) antigens by serological²⁴ and biochemical²⁵ studies. If this is the case, the notion of allelic genes in its simplest form is not applicable for the genes controlling the K - and D -region products.

Considerable physicochemical heterogeneity of K - and D -region products had been previously detected by two-dimensional electrophoresis^{26,27}. These data were, however, never connected with the antigenic diversity of H-2 molecules. The analysis of the antigenic heterogeneity of K - and D -region products may contribute to the understanding of the structure of these regions and of their function in cell-mediated immunity.

We thank Ms M. A. J. M. Treur-Mulder and A. E. Bezemer for technical assistance and Ms C. E. van Vemde for preparation of the manuscript.

Received 19 May 1980; accepted 8 January 1981.

1. Snell, G. D., Cherry, M. & Démant, P. *Transplant. Proc.* **3**, 176–179 (1971).
2. Klein, J. & Shreffler, D. C. *Transplant. Rev.* **6**, 3–29 (1971).
3. Doherty, P. C. & Zinkernagel, R. M. *J. exp. Med.* **141**, 502–507 (1975).
4. Blanden, R. V. *et al. Nature* **254**, 269–270 (1975).
5. Démant, P. *Transplant. Rev.* **15**, 162–200 (1973).
6. Snell, G. D., Cherry, M. & Démant, P. *Transplant. Rev.* **15**, 3–26 (1973).

7. Hauptfeld, V. & Klein, J. *J. exp. Med.* **142**, 288–298 (1975).
8. Lemonnier, F., Neauport-Sautes, C., Kourilsky, F. M. & Démant, P. *Immunogenetics* **2**, 517–526 (1975).
9. Démant, P. *et al. J. Immunogenet.* **2**, 263–271 (1975).
10. Démant, P. & Neauport-Sautes, C. *Immunogenetics* **7**, 295–311 (1978).
11. Iványi, D. & Démant, P. *Immunogenetics* **8**, 539–550 (1979).
12. Schwartz, B. D., Kato, K., Cullen, S. E. & Nathenson, S. G. *Biochemistry* **12**, 2157–2164 (1973).
13. Silver, J. & Hood, L. *Nature* **249**, 764–765 (1974).
14. Rask, L., Lindblom, J. B. & Peterson, P. A. *Nature* **249**, 833–844 (1974).
15. Hansen, T. H., Cullen, S. E. & Sachs, D. H. *J. exp. Med.* **145**, 438–442 (1977).
16. Neauport-Sautes, C., Morello, D., Freed, H., Nathenson, S. G. & Démant, P. *Eur. J. Immun.* **8**, 511–515 (1977).
17. Hansen, T. H. & Sachs, D. H. *J. Immun.* **121**, 1469–1472 (1978).
18. Sears, D. W. & Polizzi, C. M. *Immunogenetics* **10**, 67–82 (1980).
19. Klein, J. *Nature* **229**, 635–637 (1971).
20. Morello, D., Neauport-Sautes, C. & Démant, P. *Immunogenetics* **4**, 349–364 (1977).
21. Snoek, M., Iványi, D., Nusse, R. & Démant, P. *Immunogenetics* **8**, 109–125 (1979).
22. Senik, A. & Neauport-Sautes, C. *J. Immun.* **122**, 1461–1467 (1979).
23. Neauport-Sautes, C., Joskowicz, M. & Démant, P. *Immunogenetics* **6**, 513–527 (1978).
24. Bodmer, W. F. in *Symp. Genet. mammal. Cell Bar Harbor 1979* (in the press).
25. Orr, H. T., Lancet, D., Robb, R. J., de Castro, J. A. L. & Strominger, J. L. *Nature* **282**, 266–270 (1979).
26. Jones, P. J. *J. exp. Med.* **146**, 1261–1279 (1977).
27. Krakauer, T., Hansen, T. H., Camerini-Otero, R. D. & Sachs, D. H. *J. Immun.* **124**, 2149–2156 (1980).
28. van Nie, R. & de Moes, J. *Int. J. Cancer* **20**, 588–594 (1977).
29. Iványi, D. & Démant, P. *Immunogenetics* (in the press).

³H-baclofen and ³H-GABA bind to bicuculline-insensitive GABA_B sites in rat brain

D. R. Hill & N. G. Bowery

Department of Pharmacology, St Thomas's Hospital Medical School, London SE1 7EH, UK

The presence of a novel receptor for the neurotransmitter γ -aminobutyric acid (GABA) on peripheral autonomic nerve terminals and in mammalian brain slices has been described recently^{1–4}. This receptor differs from the classical GABA site as it is unaffected by recognized GABA antagonists such as bicuculline and is not sensitive to the majority of accepted GABA-mimetics such as 3-aminopropanesulphonic acid (3-APS) or isoguvacine. We propose to designate the classical site as the GABA_A and the novel site as the GABA_B receptor. The β -*p*-chlorophenyl derivative of GABA, baclofen, is stereospecifically active at the GABA_B site whereas it is devoid of activity at the classical GABA_A site^{2,5–9}. We now report that high-affinity saturable binding of ³H-baclofen and ³H-GABA to the GABA_B site can be detected in fragments of crude synaptic membranes prepared from rat brain. The results support the concept of a novel GABA receptor within the mammalian brain and show that GABA and baclofen can compete for the same recognition site.

The procedure adopted for this study was essentially that used by others for the detection of GABA_A sites with one notable difference—the inclusion of divalent cations in the incubation medium. Our studies have indicated that saturable high-affinity binding of ³H-baclofen is detectable in the presence of either Ca²⁺ or Mg²⁺ ions. Crude synaptic membranes were prepared by the method described by Zukin *et al.*¹⁰ and although saturable binding of ³H-baclofen (β -(4-chloro-3-³H(*N*))phenyl- γ -aminobutyric acid, 8.8 Ci mmol⁻¹) could be detected in freshly prepared and washed membranes, frozen and thawed material was routinely used. Pellets were allowed to thaw for 15 min at room temperature before suspension in Krebs–Henseleit solution (KHS), or where indicated, Tris-HCl buffer (50 mM, pH 7.4) plus CaCl₂ (2.5 mM). The tissue was maintained at 20 °C for 45 min before washing four times in fresh medium. Assays were terminated by centrifugation at 7,000g for 10 min. The specific saturable portion of binding was that displaced by an excess (100 μ M) of ($\pm$)baclofen, (–)baclofen or GABA.

In the presence of Tris-HCl buffer alone, only $5.5 \pm 0.64\%$ (mean $\pm$ s.e.m., $n = 4$ in triplicate, $P < 0.01$ paired comparison *t*-test) of the total ³H-baclofen bound could be displaced by

unlabelled ligand. In Tris-citrate buffer (50 mM, pH 7.4) no displacement could be detected (5–30 min incubation at room temperature; 5–20 nM ³H baclofen). By contrast, when KHS was used instead of Tris buffer, $21 \pm 1.1\%$ (20 nM ³H-baclofen, $n = 28$ in triplicate, equivalent to $3,712 \pm 234$ d.p.m. per mg protein) of the total could be displaced by excess unlabelled GABA, ($\pm$)baclofen or (–)baclofen (100 μ M in each case). The maximal displacement produced by each of these was the same and the extent of displacement was dose dependent.

To determine which of the ionic constituents of KHS was responsible for the emergence of the saturable component the assay was performed in Tris-HCl buffer to which one of the individual components was added. Each salt was added at the concentration present in KHS. The addition of either CaCl₂ (2.5 mM) or MgSO₄ (1.2 mM) to Tris-HCl buffer resulted in the appearance of saturable binding comparable with that observed in KHS. The addition of KCl (5 mM) or NaCl (143 mM) failed to increase the small degree of saturable binding observed in Tris-HCl alone. Thus K⁺, Na⁺ and Cl[–] ions do not seem to be

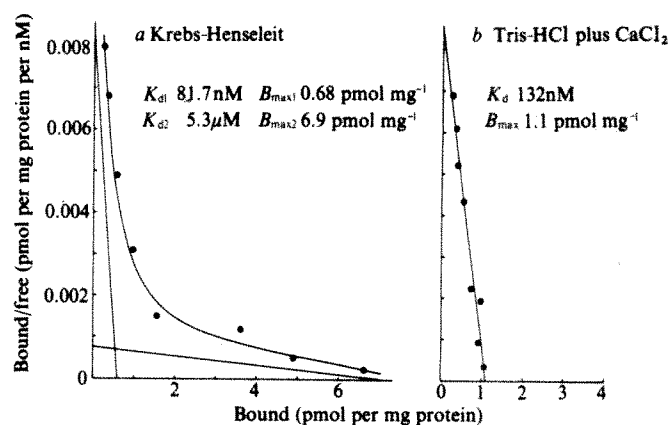


Fig. 1 Analysis of the binding of ($\pm$)³H-baclofen to rat brain synaptosomal membranes. Crude synaptic membranes were prepared by homogenization of whole rat brain in 10 vol ice-cold 0.32 M sucrose. The homogenate was centrifuged at 1,000g for 10 min and the supernatant collected and recentrifuged at 20,000g for 20 min. The pellet obtained from this second centrifugation was dispersed in distilled water (volume equal to original sucrose solution) and centrifuged at 8,000g for 20 min. The supernatant together with the buffy layer on the pellet was then centrifuged at 50,000g for 20 min. The resulting pellet was resuspended in distilled water and again centrifuged at 50,000g for 20 min. The final pellet was stored at –15 °C. For the assay in *a*, pellets which had been frozen for at least 16 h were allowed to thaw for 20 min at room temperature before resuspension in KHS. The pellet obtained from the equivalent of one rat brain was resuspended in 10 ml KHS. The suspension was incubated for 45 min at 20 °C before centrifugation at 7,000g for 10 min. This washing procedure was repeated three more times allowing 15 min incubation with each addition of KHS. The final pellet was resuspended in KHS (~1 mg protein per 0.8 ml KHS) for the assay. Protein concentrations were determined by the method of Lowry *et al.*²⁰. To each 0.8-ml aliquot of membrane suspension 0.1 ml KHS containing various concentrations of unlabelled ($\pm$)baclofen and 0.1 ml KHS containing a fixed concentration of ($\pm$)³H baclofen was added. The range of final concentrations of unlabelled baclofen was 10 nM–100 μ M. The final ³H-baclofen concentration was 20 nM. The mixture was incubated for 10 min at 20 °C and then centrifuged at 7,000g for 10 min. The supernatant was aspirated off and the pellet blotted dry before the addition of 100 μ l Soluene. The tritium level of each pellet was determined by liquid scintillation spectrometry. As the total amount bound was <3% of the radiolabel in the incubation medium, the 'free' concentration was taken as equivalent to the total tritium concentration. The same procedure was adopted in *b* as in *a*, except that Tris-HCl solution containing 2.5 mM CaCl₂ (instead of KHS) was used throughout the assay. In *a*, data derive from four separate experiments performed in triplicate at each concentration, and in *b* from a single experiment of triplicate determinations at each concentration. This experiment was repeated twice more yielding the same results.

Table 1 Comparative inhibitory activities of GABA analogues on specific ^3H -baclofen binding to synaptosomal membranes, evoked ^3H -noradrenaline release from rat isolated atria and twitch responses of mouse isolated vas deferens

Analogue	^3H -baclofen binding		Atrial ^3H -noradrenaline release ³		Vas deferens twitch ³	
	IC ₅₀ (μM)	Potency (GABA = 1)	IC ₅₀ (μM)	Potency (GABA = 1)	IC ₅₀ (μM)	Potency (GABA = 1)
GABA	0.04	1	4.2	1	3.1	1
(-) Baclofen	0.04	1	4.2	1	2.7	1.15
(±) Baclofen	0.18	0.22	4.5	0.93	2.8	1.10
(+) Baclofen	33	0.001	>500	<0.01	>100	<0.03
(±) β - <i>o</i> -chlorophenyl GABA	0.74	0.054	35	0.12	7.8	0.4
(±) β - <i>m</i> -chlorophenyl GABA	18	0.002	140	0.03	6.2	0.5
(±) β - <i>p</i> -fluorophenyl GABA	1.6	0.03	21	0.2	6.2	0.5
(±) β -hydroxy GABA	1.6	0.03	140	0.03	56	0.055
(±) β -chloro GABA	11	0.004	>500	<0.01	>300	<0.01
Muscimol	12	0.003	280	0.015	26	0.12
3-aminopropanesulphonic acid	11	0.004		NA		NA
(±) β -naphthyl GABA	300	0.0001		NA		NA
THIP	700	0.00005	600	0.007		NA
(±) β - <i>p</i> -isopropyl phenyl GABA		NA	>500	<0.01	15.5	0.2
Isoguvacine		NA		NA		NA
Taurine		NA		NA		NA
β -alanine		NA		NA		NA
Glycine		NA		NA		NA
2,4-diaminobutyric acid		NA		NA		NA
γ -hydroxybutyric acid		NA		NA	220	0.014
Nipecotic acid		NA		NA		NA
Isonipecotic acid		NA		NA		NA
(±) <i>cis</i> -3-aminocyclohexane carboxylic acid		NA		NA		NA
(±) β -aminobutyric acid		NA		NA		NA
Guanidinopropionic acid		NA		NA		NA
Imidazoleacetic acid		NA		NA		NA
(+) Bicuculline methobromide		NA		NA		NA
Picrotoxin		NA		NA		NA

NA indicates potency is <0.00005 binding; <0.001 atrium; <0.003 vas deferens. THIP, 4,5,6,7-tetrahydroisoxazolo[5,4-C]pyridine-3-ol.

essential for saturable ^3H -baclofen binding. The dependence on divalent cations could explain the difference in displaceable binding detected in Ca^{2+} -free Tris-HCl buffer compared with Tris-citrate buffer. In Tris-HCl buffer any Ca^{2+} remaining in the tissue could promote binding, whereas in Tris-citrate buffer the Ca^{2+} would be chelated and therefore unavailable for binding.

Experiments performed with varying concentrations (0–10 mM) of CaCl_2 indicated that 2.5 mM was optimal. At concentrations <1 mM the proportion of specific binding was decreased (for example, at 0.6 mM the specific portion was 30% less than at 2.5 mM). At concentrations >2.5 mM (5 and 10 mM) binding was, if anything, reduced.

Subsequent experiments to determine the characteristics of ^3H -baclofen binding were performed in KHS which contains 2.5 mM CaCl_2 and 1.2 mM MgSO_4 . This seemed appropriate as our earlier studies^{1–3} on the GABA_B site had used this solution, and as binding was apparently not Na^+ dependent, it was unlikely to be due to transport of the ligand but rather to binding at a receptor site. Moreover, the time course and lack of temperature dependence of ^3H -baclofen binding would support this. Transport of the ligand would be expected to increase with temperature but this was not observed. The amount of ^3H -baclofen specifically bound to membranes incubated at 20 °C for up to 60 min was constant. A minimum contact time with the ligand was achieved by centrifuging the incubation mixture for 5 (instead of 10) min immediately after adding the tritium. The total time taken until removal of the supernatant after centrifugation was 8 min. Even at this time saturable binding was not significantly different from that obtained after longer periods of incubation. Ten-minute incubation periods were routinely used in further experiments but clearly this was not critical.

Analysis of the binding of ^3H -baclofen in the presence of KHS revealed two saturable components. These are shown graphically in Fig. 1a using the Scatchard transformation. The kinetic parameters ($K_{d1} = 81.7 \pm 40$ nM, $B_{\text{max}1} = 0.68 \pm 0.22$ pmol per mg protein; $K_{d2} = 5.3 \pm 1.59$ μM , $B_{\text{max}2} = 6.89 \pm 0.8$ pmol per mg protein) were determined by nonlinear least-squares optimiza-

tion using a digital computer¹¹. The data were weighted according to the method of Ottaway¹². Analysis of data obtained in identical experiments but using Tris-HCl plus CaCl_2 (2.5 mM) instead of KHS revealed only a single saturable component (Fig. 1b). The parameters for this component ($K_d = 132$ nM, $B_{\text{max}} = 1.1$ pmol per mg protein, Hill coefficient = 0.997) were comparable with the high-affinity component observed in KHS. The Hill coefficient of 0.997 suggests that in Tris-HCl plus CaCl_2 binding occurs at an homogeneous site in accordance with the law of mass action.

Several unlabelled analogues have been examined for their ability to displace ^3H -baclofen from its binding sites. GABA and (–)baclofen, which were equally effective, are the most potent so far examined. (+)Baclofen was much less active (<0.001). The dose-dependency of these displacing agents is shown in Fig. 2 together with data obtained for 3-APS, muscimol, isoguvacine and bicuculline methobromide. Clearly 3-APS was only very weakly active (<0.003) and isoguvacine and bicuculline methobromide were devoid of activity. A summary of the relative potencies of the analogues is given in Table 1. There is a marked similarity between these data and those obtained previously in rat atrium and vas deferens where GABA and baclofen reduce the evoked output of transmitter. The implication is therefore that a common site exists within these systems.

Many workers have shown that baclofen is at best only weakly active in displacing ^3H -GABA from synaptic membranes suspended in Tris-citrate buffer^{13–16}. The average IC₅₀ for baclofen from these studies is >100 μM and no stereospecificity has been observed. We concur with this. However, on adding CaCl_2 (2.5 mM) to Tris-HCl buffer an additional component of ^3H -GABA binding emerges. Frozen and washed membranes were suspended in Tris-HCl buffer (50 mM, pH 7.4) containing 2.5 mM CaCl_2 and the binding assayed performed as described in Fig. 1 legend but instead of using ^3H -baclofen, ^3H -GABA (4-amino- π [2,3- ^3H]butyric acid, 10 nM, 57 Ci mmol^{–1}, Amersham) was used. Isoguvacine (final concentration 40 μM) was

present in all the assay tubes to saturate GABA_A sites. This concentration of isoguvacine produced maximal inhibition of GABA_A binding whereas it inhibited GABA_B binding by less than 10% (Fig. 2). The specific portion of GABA displaced by

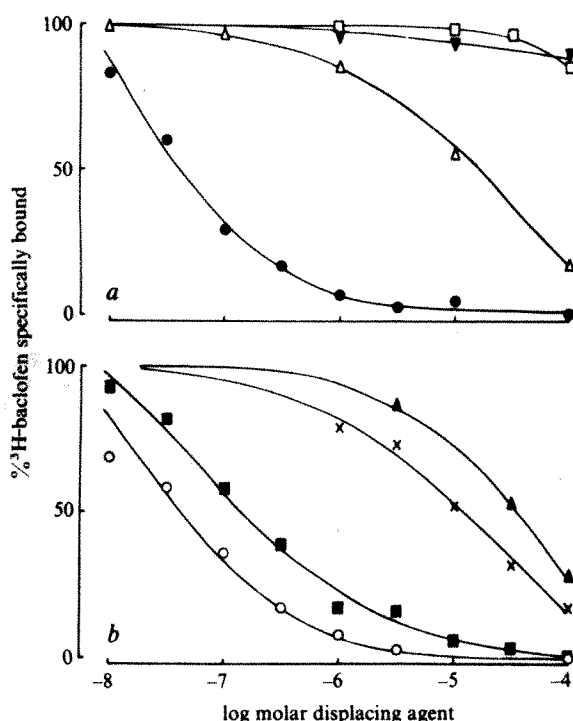


Fig. 2 Inhibition of $(\pm)^3\text{H}$ -baclofen binding by GABA-related compounds. Assays were performed according to the method described in Fig. 1 legend. Unlabelled analogues at final concentrations between 0.01 and 100 μM were added to each incubation mixture together with ^3H -baclofen (20 nM final concentration). The incubation was terminated by centrifugation at 7,000g for 10 min. Specific binding was taken as the amount of label displaced by $(\pm)$ -baclofen (100 μM). The amount displaced by individual concentrations of each analogue (log molar, abscissa) is expressed as a percentage of the total displaced by 100 μM $(\pm)$ -baclofen within the same experiment. The data have been separated for clarity. *a* Shows the effect of GABA (●), muscimol (Δ), isoguvacine (▼) and bicuculline methobromide (□). *b* Shows the effects of $(-)$ -baclofen (○), $(\pm)$ -baclofen (■), 3-aminopropanesulphonic acid (×) and $(+)$ -baclofen (▲). Each point is the mean of three experiments performed in triplicate. Standard error values were in all cases $<8\%$, and have been omitted for clarity. More than 95% of the tritium bound to the synaptosomal pellets corresponded to authentic ^3H -baclofen as determined by TLC in two systems. Six pellets (after incubation for 30 min in ^3H -baclofen) were resuspended in 1 ml 10% trichloroacetic acid and allowed to stand at room temperature for 1 h before centrifugation at 8,500g for 5 min. This procedure extracted $>86\%$ of the radioactivity of the supernatant. Aliquots (100 μl) of the supernatant were then streaked onto two silica gel plates together with markers comprising 1 pmol ^3H -baclofen, 20 μg $(\pm)$ -baclofen and 20 μg GABA. All marker solutions were prepared in 10% trichloroacetic acid. The plates were run in butanol/acetic acid/water (120:30:50) or ethanol/ammonia(0.880)/water (180:10:10). Unlabelled baclofen and GABA were visualized with 0.2% 1,2-naphthoquinone-4-sulphonic acid in 5% sodium carbonate solution and the position of labelled materials was determined by scraping sequential portions of the gel into 1 ml water before the addition of scintillation fluid. The major (*in vivo*) metabolite of baclofen, γ -hydroxy- β -chlorophenyl butyric acid was co-chromatographed on separate plates impregnated with a fluorescent indicator (Merck Keisegel 60 F₂₅₄). GABA and baclofen were also applied to these plates as markers. The metabolite and baclofen were visualized by UV illumination and GABA and baclofen again by staining. None of the pellet radioactivity corresponded to this metabolite. The experiment was repeated twice more and yielded the same results.

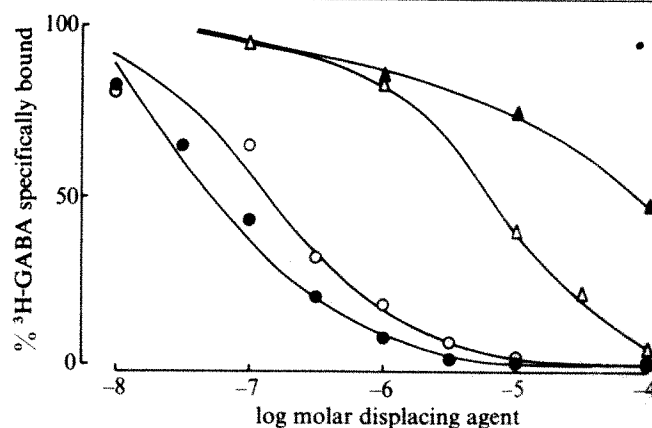


Fig. 3 Illustration of baclofen-sensitive GABA binding sites on rat brain synaptic membranes using ^3H -GABA. Membranes were prepared as described in Fig. 1 legend. Thawed pellets were similarly washed four times but instead of Krebs-Henseleit solution, Tris-HCl buffer (50 mM, pH 7.4) containing 2.5 mM CaCl₂ was used. The binding assay was also performed in a similar manner to the ^3H -baclofen assay except ^3H -GABA (10 nM) was used as the ligand. Tris-HCl plus Ca²⁺ solution was used as the incubation solution to prevent any Na⁺-dependent binding. Isoguvacine HCl (40 μM) was also present in the final incubation mixture in all samples to suppress Ca²⁺-independent GABA binding to GABA_A sites. At this concentration, isoguvacine is without effect on ^3H -baclofen binding (Fig. 2) whilst maximally displacing ^3H -GABA binding at GABA_A sites. Specific binding of ^3H -GABA was that fraction ($38 \pm 2\%$) displaced by unlabelled $(-)$ -baclofen (100 μM). Displacement by other analogues is expressed as a percentage of that produced by $(-)$ -baclofen (100 μM) in the same experiment. Each point is the mean of three determinations performed in triplicate. Standard errors were $<5\%$. ●, GABA; ○, $(-)$ -baclofen; Δ, muscimol; ▲, $(+)$ -baclofen.

100 μM $(-)$ -baclofen represented $18,751 \pm 1,511$ d.p.m. per mg protein which was equivalent to $38 \pm 2\%$ of binding ($n = 13$ in triplicate). This Ca²⁺-dependent GABA binding was saturable and Scatchard analysis indicated a single component ($K_d = 77$ nM, $B_{\text{max}} = 1.22$ pmol per mg protein, Hill coefficient = 1.06 mean of six experiments). The structural specificity of compounds which inhibited binding was comparable with that observed for displacement of ^3H -baclofen (Fig. 3). The 'cross-displacement' between GABA and baclofen strongly suggests that they have affinity for a common recognition site which differs from the 'classical' GABA_A site.

We believe that this is the first binding system to be described which is dependent on the presence of divalent cations. Enhancement of the binding of opiate agonists and glutamic acid by divalent cations has been demonstrated. However, the activation of ^3H -glutamate binding by calcium is markedly temperature dependent and persists after removal of the cation^{17,18}. This does not occur at the GABA_B site. The activation of Na⁺-independent ^3H -dihydromorphine binding by calcium is very weak and it is unlikely that calcium regulates opiate receptor binding¹⁹. By contrast, calcium or magnesium seem to be necessary to detect binding at the GABA_B site.

Thus, we have shown that tritiated baclofen and GABA can bind in a saturable manner to a receptor site on mammalian synaptic membranes, and that GABA and $(-)$ -baclofen have equal affinity for this binding site. Pharmacologically the site seems to be identical to that previously described on peripheral nerve terminals¹⁻³ (Table 1) and in mammalian brain slices⁴. Bicuculline-insensitive GABA sites seem to be present on nerve terminals and, when activated, reduce evoked neurotransmitter release. Therefore, the GABA_B binding site described in this study is probably present on nerve terminals.

We thank Ciba-Geigy Basle for supplies of ^3H -baclofen and the enantiomers of baclofen, and Ciba-Geigy Horsham for financial assistance. Isoguvacine, bicuculline methobromide and

THIP were given by Drs J. F. Collins and P. Krogsgaard-Larsen. We also acknowledge the assistance of Dr R. White, who performed the computer analysis, and Mrs J. Andrews, who typed the manuscript. D.R.H. is an SRC Scholar.

Received 3 November 1980; accepted 13 January 1981.

1. Bowery, N. G. & Hudson, A. L. *Br. J. Pharmac.* **66**, 108P (1979).
2. Bowery, N. G. *et al. Br. J. Pharmac.* **67**, 444-445P (1979).
3. Bowery, N. G. *et al. Eur. J. Pharmac.* (in the press).
4. Bowery, N. G. *et al. Nature* **283**, 92-94 (1980).
5. Curtis, D. R., Game, C. J. A., Johnston, G. A. R. & McCulloch, R. M. *Brain Res.* **70**, 493-499 (1974).
6. Davies, J. & Watkins, J. C. *Brain Res.* **70**, 501-505 (1974).
7. Olpe, H. R., Koella, W. P., Wolf, P. & Haas, H. L. *Brain Res.* **134**, 577-580 (1977).
8. Ault, B. & Evans, R. H. *J. Physiol., Lond.* **284**, 131P (1978).
9. Feltz, P., Deschenes, M. & Desarmenien, M. in *Iontophoresis and Transmitter Mechanisms of the Mammalian Central Nervous System* (eds Ryall, R. & Kelly, J. S.) 264-266 (Elsevier, Amsterdam, 1978).
10. Zukin, S. R., Young, A. B. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4802-4807 (1974).
11. Powell, M. J. D. *Harwell Report AERE-R5947* (HMSO, London, 1968).
12. Ottaway, J. H. *Biochem. J.* **134**, 729-736 (1973).
13. Galli, A., Zilletti, L., Scotton, M., Adembi, G. & Giotti, A. *J. Neurochem.* **32**, 1123-1125 (1979).
14. Olsen, R. W., Greenlee, D., Van Ness, P. & Ticku, M. K. in *Amino Acids as Chemical Transmitters* (ed. Fonnum, F.) (Plenum, New York, 1978).
15. Waddington, J. L. & Cross, A. J. *Naumyn-Schmiedeberg's Archs Pharmac.* **306**, 275-280 (1979).
16. Horng, J. S. & Wong, D. T. *J. Neurochem.* **382**, 1379-1386 (1979).
17. Baudry, M. & Lynch, G. *Nature* **282**, 748-750 (1979).
18. Baudry, M. & Lynch, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2298-2302 (1980).
19. Pasternak, G. W., Snowman, A. M. & Snyder, S. H. *Molec. Pharmac.* **11**, 735-744 (1975).
20. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).

An oestrogen receptor activator protein in rat uterine cytosol

T. N. R. V. Thampan & James H. Clark

Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA

Earlier reports of oestrogen receptor binding to DNA¹⁻⁴ indicated that a protein which does not bind oestrogen might be involved in the conversion of the 4S oestradiol receptor to the 5S form but no direct evidence for the existence of this protein was presented. This receptor transformation or activation step is thought to be a requirement for events in the nucleus. We considered it likely that such a protein might be similar to the cytoplasmic proteins which stimulate eukaryotic RNA polymerases⁵⁻¹¹. These are basic proteins which are not adsorbed to DEAE cellulose and have a sedimentation coefficient of ~3S. To examine this possibility, we fractionated unlabelled uterine cytosol by DEAE-cellulose chromatography. Unlabelled cytosol was used to avoid the possibility that such a factor might be retained on the column as a component of the receptor-oestradiol complex. Here we present evidence of a protein present in uterine cytosol which forms a complex with the 4S form of the oestrogen receptor to produce a 5S-activated form. This activated form binds to uterine nuclei, DNA-cellulose and native calf thymus DNA. The cytosol protein, which we call receptor activation factor, does not bind oestradiol, has a sedimentation coefficient of ~3S and does not bind to DEAE-cellulose.

Adult female rats of the Holtzman strain were used 2 weeks after ovariectomy. All experimental procedures were carried out at 4 °C. Two uteri were homogenized in 10 ml TETK buffer (0.01 M Tris-HCl, 0.001 M EDTA, 12 mM thioglycerol and 0.075 M KCl, pH 7.8) and centrifuged at 10,000 r.p.m. for 15 min. The supernatant was centrifuged at 45,000 r.p.m. for 2 h and the cytosol fraction recovered. DEAE-cellulose (Whatman DE 52) was equilibrated with TETK buffer and washed extensively with the buffer before chromatography.

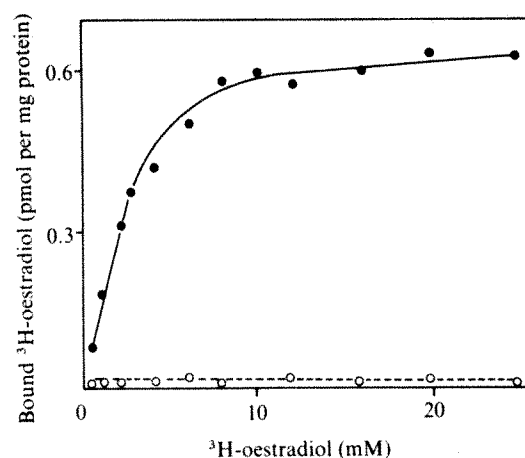


Fig. 1 Oestradiol binding to uterine cytosol fractions as a function of hormone concentration. Fraction I (○) and fraction II (●) of uterine cytosol were prepared as described in the text. Aliquots of each fraction were incubated with varying concentrations of ³H-oestradiol (specific activity 90 Ci mmol⁻¹) in the presence or absence of 100-fold excess of DES in a total volume of 350 μl at 30 °C for 30 min. After incubation, the samples were cooled in ice and 500 μl of a 60% suspension of HAP added. The samples were vortexed, left in ice for 10 min and centrifuged at 3,000 r.p.m. for 5 min. The HAP pellet was washed three times with 0.01 M Tris-HCl, 0.001 M EDTA, pH 7.4. The pellet was then extracted with 1 ml ethanol and the radioactivity assayed using ACS (Amersham). Protein was estimated according to Lowry *et al.*¹⁵.

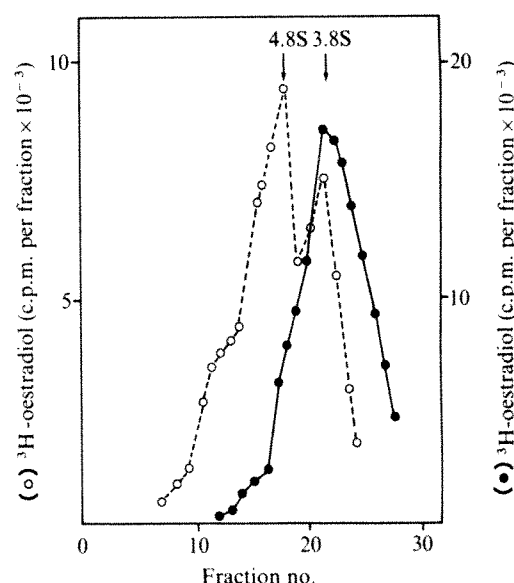


Fig. 2 Effect of fraction I on the sedimentation behaviour of the fraction II-oestradiol complex. Fraction II (45 μg of protein which contained the cytosol oestrogen receptor) was incubated with 10 nM ³H-oestradiol in the presence or absence of fraction I (40 μg protein) at 30 °C for 30 min. The labelled materials were subjected to density-gradient centrifugation on 5-20% linear sucrose gradients (in TETK buffer containing 0.3 M KCl) at 45,000 r.p.m. for 16 h. Parallel gradients containing 1 μM DES in addition to ³H-oestradiol were run to compute the specific hormone binding. Human γ globulin and ovalbumin were used as sedimentation markers. ●, Fraction II + TETK buffer; ○, fraction I + fraction II.

Cytosol was applied to the DEAE-cellulose column (2 × 6 cm) and allowed to enter the gel, which was washed with TETK buffer after 1 h. The fractions that did not adsorb to the DEAE-cellulose were recovered (fraction I). The column was then washed with TETK buffer containing 0.3 M KCl and the eluted material collected (fraction II). Solid (NH₄)₂SO₄ was added to the two fractions to 75% saturation to precipitate the

macromolecules and the precipitate was dissolved in the required volume of TETK buffer (2 and 4 ml for fractions I and II respectively).

Concentration-dependent binding of ^3H -oestradiol to the macromolecules in fractions I and II was studied using the hydroxyapatite (HAP) assay method as described in Fig. 1 legend and in Clark and Peck¹². Fraction II showed a saturation curve typical of hormone-receptor preparations, but no significant binding of oestradiol was found in fraction I (Fig. 1). The effects of the addition of fraction I on hormone-binding properties of fraction II were examined. Fraction II-oestradiol complex sedimented as a 4S (3.8S) peak on sucrose density gradients in the absence of fraction I. However, when equivalent protein concentrations of fraction I were mixed with fraction II two peaks of bound hormone were observed, one sedimenting at

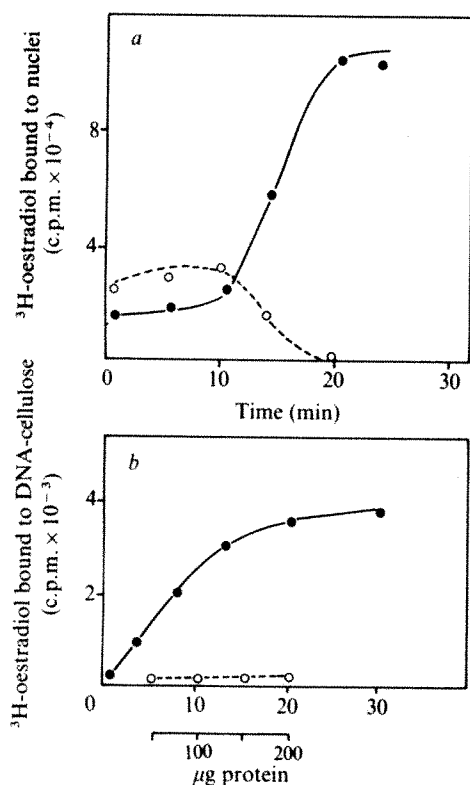


Fig. 3 *a*, Effect of addition of fraction I on the binding of fraction II-oestradiol complex to isolated uterine nuclei. Nuclei (25 μg DNA) were incubated with 10 nM ^3H -oestradiol $\pm$ 1 μM DES and fraction II (40 μg protein) either in the presence (●) or absence (○) of fraction I (20 μg protein) at 30 °C for various time periods. At the end of the incubation the tubes were centrifuged and the pelleted nuclei washed three times with TETK buffer. The ^3H -oestradiol bound to the nuclei was extracted with ethanol and the bound hormone calculated. DNA was measured according to the procedure of Burton¹⁶. *b*, Binding of fraction II-oestradiol complex to DNA-cellulose as a function of the concentration of fraction I proteins. DNA-cellulose (10 μg DNA) was incubated at 30 °C for 30 min with 10 nM ^3H -oestradiol, fraction II (50 μg protein) and various concentrations of fraction I (●). Similar experiments were done with various levels of fraction II in the absence of fraction I (○). Parallel incubations with 1 μM DES were also done to establish specific binding. After incubation, the samples were cooled in ice and 1 ml of cold TETK buffer added. The DNA-cellulose was pelleted after centrifugation for 5 min at 4 °C (3,000 r.p.m.) and washed three times using 1 ml volumes of TETK buffer. The radioactivity bound to DNA-cellulose was extracted with ethanol. The counts associated with the DNA-cellulose in the presence of 1 μM DES were subtracted from the total counts to assess the specific binding of the receptor to DNA-cellulose.

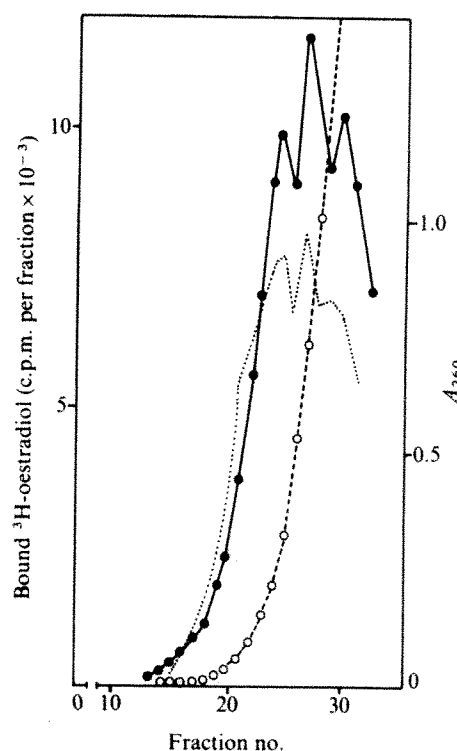


Fig. 4 Effect of fraction I on the binding of fraction II to native calf thymus DNA. 20 μg DNA was incubated with fraction II (40 μg protein) and 10 nM ^3H -oestradiol in the presence (●) or absence (○) of fraction I (30 μg protein) at 30 °C for 30 min. Centrifugation on linear 5–20% sucrose gradients (in TETK buffer) was done at 45,000 r.p.m. for 2 h. The dotted line represents absorbance at 260 nm.

5S (4.8S) and the other at 4S (3.8S) (Fig. 2). In the presence of fraction I, radioactivity associated with the 4S fraction was reduced to less than 50% of that detected in its absence. The radioactivity that was lost from the 4S appeared in the 5S peak. In all these cases the protein fractions were incubated with 10 nM ^3H -oestradiol in the presence or absence of 1 μM diethylstilboestrol (DES) at 30 °C for 30 min before density-gradient centrifugation and the nonspecific binding was subtracted from the total to determine specific binding.

Time-dependent binding of fraction II-oestradiol complex to isolated uterine nuclei was investigated. Uterine nuclei were purified as follows: uteri were homogenized in TMKC-sucrose buffer (10 mM Tris-HCl, pH 7.5, 5 mM MgCl_2 , 25 mM KCl, 1 mM CaCl_2 and 0.5 M sucrose) using a ground glass homogenizer and centrifuged at 2,000 r.p.m. (Beckman JS7.5 rotor) for 10 min. The pellet was gently homogenized in the same buffer containing 0.25% Triton X-100 and 0.25% NP40, filtered through two layers of organza and centrifuged. The pellet was resuspended in the buffer, filtered through organza and centrifuged twice to remove the detergent. The washed pellet was resuspended in TMKC buffer containing 1.9 M sucrose and centrifuged at 12,000 r.p.m. (Beckman JS-13 rotor) for 30 min. The final nuclear pellet was resuspended in TETK buffer. Incubation was carried out immediately after the preparation of the nuclei. Fraction II, in the absence of fraction I, was incubated with nuclei and showed initial binding which was then reduced to insignificant levels after 20 min (Fig. 3*a*). In contrast, the nucleus-bound hormone showed a sharp increase when the nuclei were incubated in the presence of fractions I and II. An increase in oestradiol binding to DNA-cellulose¹³, dependent on the concentration of fraction I, was observed when DNA-cellulose suspensions were mixed with a constant

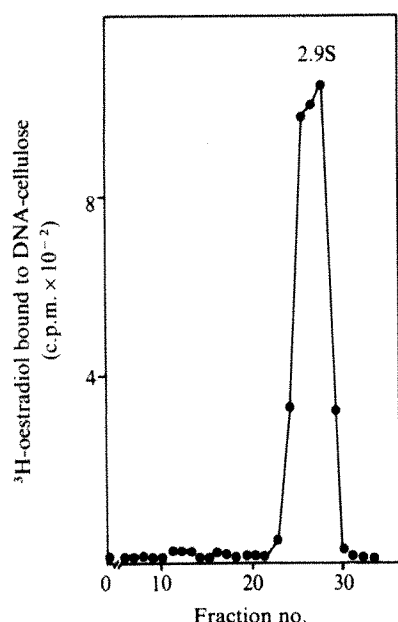


Fig. 5 Sedimentation behaviour of the active factor in fraction I that stimulates the binding of fraction II-oestradiol complex to DNA-cellulose. An aliquot of fraction I was centrifuged on a 5–20% linear sucrose gradient (in 0.3 M KCl-TETK buffer) at 45,000 r.p.m. for 16 h. The fractions collected from the punctured tube were assayed for the presence of the receptor activation factor by incubating with DNA-cellulose (5 µg DNA), fraction II (10 µg protein) and 10 nM ³H-oestradiol ± 1 µM DES at 30 °C for 30 min. The specific binding was calculated as described in Fig. 3 legend.

amount of fraction II protein and varying doses of fraction I (Fig. 3b). There was no specific binding of oestradiol to DNA-cellulose in the absence of fraction I, irrespective of large increases in the amount of fraction II protein added. Pretreatment of fraction I with trypsin destroyed its ability to cause binding of fraction II to DNA-cellulose (data not shown).

Native calf thymus DNA was incubated with fraction II and ³H-oestradiol in the presence or absence of fraction I (Fig. 4) and the samples were layered over 5–20% sucrose density gradients and centrifuged at 45,000 r.p.m. for 2 h. Gradient fractions were examined for the binding of oestradiol-fraction II complex to DNA—no binding was observed when the incubation was carried out in the absence of fraction I but when fraction I was mixed with fraction II the labelled complex co-sedimented with the DNA.

The ability of fraction I to effect the binding of fraction II-oestradiol complex to DNA-cellulose was the basis of an assay for the detection of the active factor in fraction I. An aliquot of fraction I was subjected to density-gradient centrifugation on 5–20% sucrose gradients at 45,000 r.p.m. for 16 h. The fractions were individually incubated with known amounts of fraction II protein, DNA-cellulose and 10 nM ³H-oestradiol either in the presence or absence of 1 µM DES. A macromolecular fraction that sedimented at 3S (2.9S) was shown to stimulate oestradiol-fraction II binding to DNA-cellulose (Fig. 5)—we call this protein the receptor activation factor.

These data show that a cytosol protein which does not bind oestradiol is involved in the activation of the oestrogen receptor. This activation factor causes a 4 to 5S shift in the sedimentation coefficient of the receptor and allows it to bind to DNA and uterine nuclei. It is possible that the basic protein identified by Bresciani *et al.*¹⁴ as an intranuclear acceptor for oestrogen receptors is similar to the receptor activation factor, which is tightly bound to chromatin following nuclear translocation. There is an obvious analogy between the activation factor and factors which stimulate RNA polymerase activity in other

eukaryotic systems^{5–11}. As suggested by Thrower *et al.*⁴, it is possible that the function of the receptor is to translocate this factor to the nucleus, where it stimulates transcriptional events.

This work was supported by a WHO research training grant to T.N.R.V.T. and NIH grant HD 08436 to J.H.C.

Received 28 July 1980; accepted 9 January 1981.

1. Yamamoto, K. R. *J. biol. Chem.* **249**, 7068–7075 (1974).
2. Notides, A. C. & Neilson, S. *J. biol. Chem.* **249**, 1866–1873 (1974).
3. Yamamoto, K. R. & Alberts, B. M. *Cell* **4**, 301–310 (1975).
4. Thrower, S., Hall, C., Lim, L. & Davison, A. N. *Biochem. J.* **160**, 270–280 (1976).
5. Seifart, K. H. *Cold Spring Harb. Symp. quant. Biol.* **35**, 719–725 (1970).
6. Stein, H. & Hausen, P. *Cold Spring Harb. Symp. quant. Biol.* **35**, 709–717 (1970).
7. Natori, S. *J. Biochem.* **72**, 1291–1294 (1972).
8. Lentfer, D. & Lezius, G. *Eur. J. Biochem.* **30**, 278–284 (1972).
9. Natori, S., Takeuchi, K., Takahashi, K. & Mizuno, D. *J. Biochem.* **73**, 879–888 (1973).
10. Lee, S. C. & Dahmus, M. E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1383–1387 (1973).
11. Seifart, K. H., Juhasz, P. P. & Benecke, B. J. *Eur. J. Biochem.* **33**, 181–191 (1973).
12. Clark, J. H. & Peck, E. J. Jr *Female Sex Steroids: Receptors and Function*, 4–36 (Springer, Berlin, 1979).
13. Alberts, B. M. & Herrick, G. *Meth. Enzym.* **21**, 198–217 (1971).
14. Bresciani, F., Puca, G. A., Nola, E. & Sica, V. in *Control of Proliferation of Animal Cells* (eds Clarkson, B. & Baserga, R.) 67–83 (Cold Spring Harbor Laboratory, New York, 1974).
15. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
16. Burton, K. *Biochem. J.* **62**, 315–323 (1956).

***onc* sequences (*v-fes*) of Snyder-Theilen feline sarcoma virus are derived from noncontiguous regions of a cat cellular gene (*c-fes*)**

G. Franchini*, J. Even†, C. J. Sherr† & F. Wong-Staal*

* Laboratory of Tumor Cell Biology and † Viral Pathology Section, National Cancer Institute, NIH, Bethesda, Maryland 20205, USA

Type C sarcoma viruses are genetic recombinants containing portions of replication-competent helper viruses linked to sarcoma virus-specific sequences (generically designated *onc* genes) which are thought to be required for acute fibroblast transformation^{1–4}. The *onc* elements of different avian and mammalian sarcoma viral isolates are each homologous to subsets of cellular DNA sequences which have no well-defined role in normal cells^{5–8}. Because of the lack of significant homology between helper viral genes and cellular *onc* sequences^{1–9}, the recombinational mechanisms which facilitate the formation of sarcoma viral genomes remain unclear. In Moloney murine sarcoma virus, viral *onc* (or *v-mos*) and cellular *onc* (or *c-mos*) sequences exhibit complete and uninterrupted homology as determined by heteroduplex and restriction enzyme analyses of molecularly cloned DNA⁹. By contrast, the cellular counterparts of the *onc* elements of Rous sarcoma virus (G. Cooper and R. Parker, personal communication), avian erythroblastosis virus (B. Vennstrom, personal communication), Abelson leukaemia virus (D. Baltimore, personal communication), Harvey sarcoma virus (E. Scolnick, personal communication) and simian sarcoma virus (R. Gallo, personal communication) are now known to contain intervening sequences which do not appear in the respective viral genomes. Here we report the use of the Southern blot technique¹⁰ to examine cat cellular DNA sequences (*c-fes*) homologous to the *onc* gene (*v-fes*) of Snyder-Theilen feline sarcoma virus (ST-FeSV)¹¹. We used cloned DNA 'probes' containing defined portions of the ST-FeSV genome to show that *v-fes* sequences originate from at least four noncontiguous sequences in cat cellular DNA, separated from each other by intervening sequences.

The unintegrated DNA intermediates of ST-FeSV and feline leukaemia virus (FeLV) (subgroup B) helper virus have been previously cloned in a bacteriophage λ vector and characterized in detail¹². The *v-fes* sequences of ST-FeSV are ~1.3 kilobase pairs long and include at least three cleavage sites for the restriction endonuclease *Pst*I. An additional *Pst*I site is located to the right and demarcates a 140-base fragment which includes the 3' junction of *v-fes* and FeLV-derived sequences (Fig. 1). The two larger *Pst*I fragments, each ~0.5 kilobase pairs long, we have designated *S_L* and *S_R* as they represent the left and right portions of the *v-fes* region in 5' to 3' orientation with respect to viral RNA. *S_L* and *S_R* DNA fragments were partially purified by electroelution from agarose gels and separately subcloned into the ampicillin-resistance gene of the plasmid vector pBR322. The recombinant plasmids were identified by replica-plating and colony hybridization techniques, grown in an *Escherichia coli* host, and their DNA extracted using methods described elsewhere¹³. Purified plasmids containing *S_L* or *S_R* were labelled with ³²P by nick-translation¹⁴ and hybridized by the Southern blot technique¹⁰ to restricted DNA obtained from tissues of specific pathogen-free domestic cats (Liberty Laboratories).

The restriction endonucleases *Hind*III, *Eco*RI and *Bgl*II do not recognize cleavage sites within *v-fes* sequences¹⁵ and when cat DNA was cleaved with these enzymes, single *c-fes*-containing fragments of 9.4 (*Hind*III), 13 (*Eco*RI) and 22 (*Bgl*II) kilobase pairs were detected with both radiolabelled probes (Fig. 2, lanes 1–3). The *S_L* and *S_R* portions of *c-fes* must therefore be sited at a discrete genetic locus (maximal complexity = 9.4 kilobase pairs).

If sequences homologous to the *S_L* probe were contiguous in cat cellular DNA, cleavage with *Pst*I should yield a single hybridizable DNA fragment equal in length to that of *S_L* itself (~500 bases). However, *Pst*I generated two fragments of 1.8 and 0.7 kilobase pairs which were detected with the *S_L* probe (Fig. 2, lane 6 and Fig. 3, lane 1). The combined length of these fragments (2.5 kilobase pairs) is at least 1.9 kilobase pairs longer than the *S_L* sequence itself which indicates that the *c-fes* domain related to *S_L* must contain 1.9 kilobase pairs of intervening sequences that include a *Pst*I site absent in *v-fes*. To orientate the 0.7- and 1.8-kilobase pair fragments from 5' to 3' with respect to viral RNA, we performed co-digestion analyses with enzymes that recognize single, asymmetric sites within the *v-fes* *S_L* region. Figure 1 shows that *Kpn*I and *Sac*I each cleave at opposite ends of the *S_L* region within 0.1 kilobase pairs of the *Pst*I termini. Co-digestion of cat DNA with *Kpn*I and *Sac*I yielded a 2.3-kilobase pair fragment which hybridized strongly to the *S_L* probe (Fig. 3, lane 9). The length of this fragment was

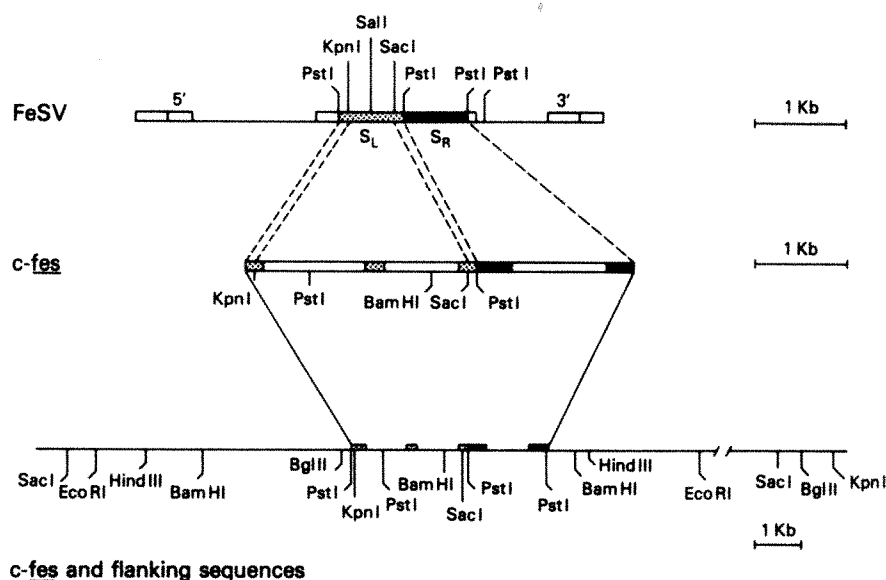
~0.2 kilobase pairs shorter than the total complexity of the *c-fes* *S_L* domain (2.5 kilobase pairs) which demonstrates that no *Kpn*I or *Sac*I sites can be detected in the region(s) containing intervening sequences. Digestion with *Kpn*I reduced the length of the 0.7-kilobase pair *Pst*I fragment by <0.1 kilobase pairs (Fig. 3, lanes 1, 2), and *Sac*I reduced the length of the 1.8-kilobase pair *Pst*I fragment by a similar extent (Fig. 3, lanes 1, 3). Thus, the 0.7-kilobase pair *Pst*I fragment is located at the 5' end of the 1.8-kilobase pair *Pst*I fragment (Fig. 1).

The *S_R* probe also hybridized to a 1.7-kilobase pair *Pst*I fragment which is about three times more complex than the *S_R* sequences themselves (Fig. 2, lane 7). No enzymes tested so far have cleaved the *S_R*-related 1.7-kilobase pair *Pst*I fragment. The simplest model places an additional 1.2 kilobase pairs of intervening sequences at the centre of the *c-fes* *S_R* domain as shown in Fig. 1. These data suggest that the complexity of the *c-fes* domain detected with both probes is 4.2 kilobase pairs (*S_R*, 1.7 and *S_L*, 2.5 kilobase pairs).

Co-digestion experiments and hybridization with the two probes allowed the sites of cleavage for *Hind*III and *Eco*RI to be positioned with respect to the defined *Sac*I and *Kpn*I sites in *c-fes* (data not shown). Each of these enzymes recognizes sites which are comparatively distant from the *c-fes* region. Other positions for *Sac*I and *Kpn*I cleavage in flanking DNA sequences were assigned by similar procedures (see Fig. 2, lanes 4, 5 and Fig. 3, lanes 7, 9 for *Sac*I and *Kpn*I data). In contrast to *Hind*III and *Eco*RI, *Bgl*II recognizes a cleavage site close to the 5' end of the *c-fes* *S_L* domain. Co-digestion with *Sac*I and *Bgl*II generates a 2.6-kilobase pair fragment detectable with the *S_L* probe (Fig. 3, lane 8), which positions the *Bgl*II site within 0.3 kilobase pairs from the left-hand *S_L* *Pst*I site (Fig. 1).

Although *v-fes* lacks sites of restriction for *Bam*HI (ref. 15), cleavage with this enzyme yielded two DNA fragments of 5.2 and 2.8 kilobase pairs which were detected with the *S_L* probe (Fig. 2, lane 8 and Fig. 3, lane 5). The 2.8-kilobase pair fragment alone was seen using the *S_R* probe (Fig. 2, lane 9) which indicates that *Bam*HI, like *Pst*I, recognizes a single site within the *c-fes* *S_L* domain. As expected, the left-hand 5.2-kilobase pair *Bam*HI fragment was cleaved by *Bgl*II to produce a 2.3-kilobase pair *S_L*-related fragment, thus positioning the *Bam*HI site (Fig. 3, lanes 5, 6). Co-digestion of cat DNA with *Bam*HI and *Pst*I was therefore expected to result in the digestion of the 1.8-kilobase pair *Pst*I fragment to two fragments of ~1.4 and 0.4 kilobase pairs. If the 1.4-kilobase pair *Pst*I–*Bam*HI fragment thus produced contained only intervening sequences, we would not expect to detect it with the *S_L* probe. Surprisingly, this fragment was readily visualized (Fig. 3, lane 4) which indicates that at least

Fig. 1 Restriction endonuclease map of cat cellular DNA sequences containing regions of homology to *v-fes*. The top line represents the linear ST-FeSV provirus from 5' to 3' with respect to viral RNA. Subcloned *Pst*I fragments representing 80% of *v-fes* (box) and labelled *S_L* and *S_R* were used to map homologous sequences in restriction fragments of cat cellular DNA. The middle line shows the sequences homologous to *S_L* and *S_R* probes with *c-fes*. The complexity and location of the central hybridizing sequence (stippled area) at the left of the *Bam*HI site remain uncertain. Sequences homologous to *v-fes* *S_R* shown in black contain at least one set of intervening sequences no longer than 1.2 kilobase pairs. The bottom diagram shows non-inclusive sites of restriction in cat DNA sequences flanking *c-fes*. Only sites which can be assigned using *S_L* and *S_R* probes are indicated.



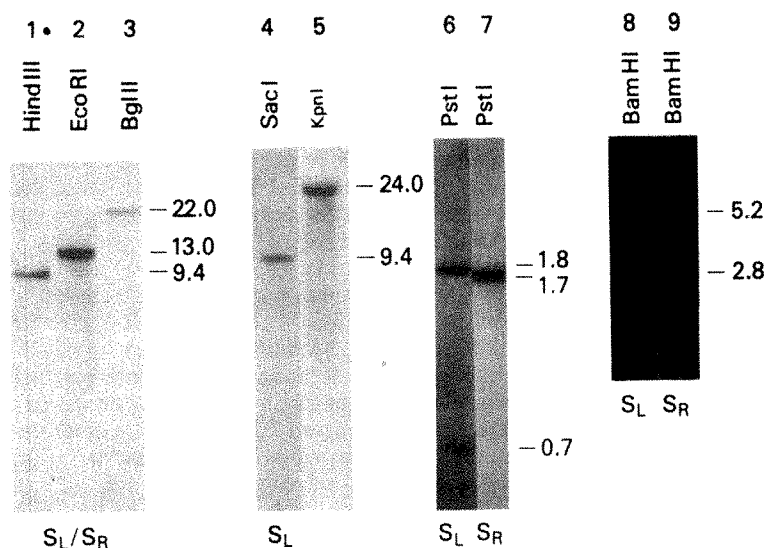


Fig. 2 Detection of sequences in cat cellular DNA related to *v-fes* S_L and S_R probes. DNA (30 μ g per lane) was digested and hybridized in conditions described by Wahl *et al.*²⁵. The sizes of fragments (in kilobase pairs, kb) were determined from the positions of restriction fragments of adenovirus type 2 DNA (Bethesda Research Laboratories) used as internal digestion standards. Lanes 1-3 show results from a 0.8% gel using the S_L probe; identical results were obtained with S_R or with a mixture of S_L and S_R probes. Lanes 4, 5, 8 and 9 show results from 1.4% gels and lanes 6 and 7 from a 1% gel. The enzymes used are indicated above the lanes and the probes are noted below.

some S_L -related sequences map within this region. The S_L -related domain of *c-fes* must therefore contain at least two noncontiguous segment of intervening sequences (Fig. 1).

The presence of a *SalI* site within S_L sequences suggested that it might be possible to localize the central hybridizing sequence more precisely within the *c-fes* S_L domain. However, co-digestion of cat DNA with *SalI* and several other enzymes failed to demonstrate a *SalI* site in *c-fes*. Although we favour the possibility that the *SalI* hexanucleotide recognition sequence is methylated, we cannot exclude the possibility that the *SalI* site is in fact absent.

Thus, cloned subgenomic DNA fragments representing ~80% of the ST-FeSV sarcoma virus-specific sequences can be used to detect a discrete *c-fes* locus in the cellular DNA of normal domestic cats. This locus (estimated complexity of S_L and S_R , 4.2 kilobase pairs) is significantly more complex than *v-fes* itself and contains at least three subsets of intervening sequences which lack homology to the viral *onc* gene. Although the presence of intervening sequences is suggested by the complexity of the *c-fes* locus alone, our positioning of the intervening sequences is based on the simple assumption that the restriction sites in *v-fes* are retained in *c-fes*. In the case of *v-fes* S_L sequences where five sites for restriction endonuclease cleavage have been assigned, at least four (and possibly all) of these seem to be retained in the *c-fes* S_L domain. Two additional

sites for *BamHI* and *PstI* also appear within the *c-fes* S_L domain and map within two different intervening sequences. Because no sites other than those for *PstI* have been assigned within the *v-fes* S_R region itself or within the *c-fes* S_R domain, we cannot determine whether the S_R domain includes one or more subsets of intervening sequences, nor can we exclude the possibility that a *PstI* site in *v-fes* has no homologue in *c-fes*. Finally, we have not positioned the sequences in cellular DNA that are homologous to the extreme 5' end (0.2-0.4 kilobase pairs) and 3' end (0.1 kilobase pairs) of *v-fes* for which specific probes are unavailable. The close proximity of the *BglII* site to the 5' end of the S_L domain raises the possibility, however, that this site maps within yet another *c-fes* intervening sequence. Thus, it seems likely that *c-fes* is indeed more complex than our minimal estimate of 4.6 kilobase pairs (S_L , S_R and unmapped terminal *v-fes* sequences).

The organization of the *c-fes* gene raises critical questions concerning the formation of feline sarcoma viral genomes. ST-FeSV encodes a polyprotein of 85,000 molecular weight^{16,17} which shows protein kinase activity^{18,19} and is specified in part by *v-fes* sequences¹⁵. Transformation-defective ST-FeSV mutants²⁰ synthesize altered polyproteins that have no kinase activity (M. Barbacid, S. K. Ruscetti, L. Donner and C. J. Sherr, unpublished data) which strongly suggests that expression of the *v-fes* encoded function is required for ST-FeSV-induced transformation. If, like the Rous sarcoma viral system²¹⁻²³, the cellular *onc* gene codes for a protein which is immunologically and functionally related to that encoded by ST-FeSV, the intervening sequences would presumably represent noncoding regions of the cellular gene (introns) which would be deleted after formation of the primary transcripts. Although formation of recombinant FeSV genomes could involve cross-over between a helper DNA provirus and *c-fes* DNA sequences, the retention of only the putative coding regions (exons) would then have to involve a process of systematic DNA deletions analogous to RNA splicing events. Alternatively, spliced *c-fes* mRNA molecules could be packaged in helper viral particles, copied by reverse transcription into DNA in another round of infection, and be subsequently recombined with proviral DNA. Whatever the mechanism, the formation of transforming sarcoma viruses would at least require the retention of the critical coding sequences of the cellular gene. As the Gardner-Arnstein (GA) strain of FeSV²⁴ has also acquired sequences homologous to *v-fes*⁸, it will be of interest to see if the differences between the ST and GA strains^{16,17} reside in the sarcoma virus-specific or helper virus-derived regions of these genomes.

We thank Dr William McClements for his advice in preparing the plasmid subclones, Dr Robert C. Gallo for helpful criticisms and support, K. Gorse for assistance with the experiments, and the many investigators for communication of the unpublished results.

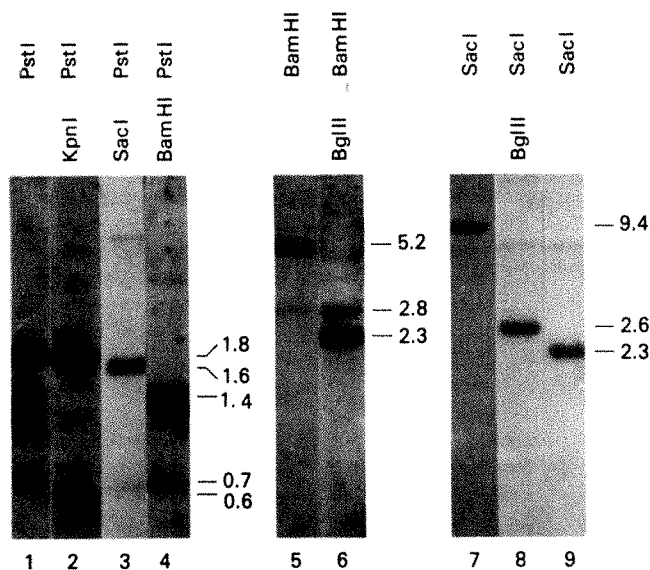


Fig. 3 Co-digestion experiments done using the S_L probe. The enzymes used are indicated above the lanes and the lengths of the fragments in kilobase pairs are shown at the sides.

Received 8 October 1980; accepted 2 January 1981.

1. Vogt, P. K. *Comp. Virol.* **9**, 341-455 (1977).
2. Hanafusa, H. *Comp. Virol.* **10**, 401-483 (1977).
3. Bishop, J. M. A. *Rev. Biochem.* **47**, 34-88 (1978).
4. Fischinger, P. J. in *Molecular Biology of RNA Tumor Viruses* (ed. Stephenson, J. R.), 163-198 (Academic, New York, 1979).
5. Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. *Nature* **260**, 170-173 (1976).
6. Scolnick, E. M., Rands, E., Williams, D. & Parks, W. P. *J. Virol.* **12**, 458-463 (1973).
7. Frankel, A. E. & Fischinger, P. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3705-3709 (1976).
8. Frankel, A. E., Gilbert, J. H., Porzig, K. J., Scolnick, E. M. & Aaronson, S. A. *J. Virol.* **30**, 821-827 (1979).
9. Oskarsson, M., McClements, W. L., Blair, D. G., Maizel, J. V. & Vande Woude, G. F. *Science* **207**, 1222-1224 (1980).
10. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
11. Snyder, S. P. & Theilen, G. H. *Nature* **221**, 1074-1075 (1969).
12. Sherr, C. J., Fedele, L. A., Oskarsson, M., Maizel, J. & Vande Woude, G. F. *J. Virol.* **34**, 200-212 (1980).
13. Bolivar, F. *et al. Gene* **2**, 95-113 (1977).
14. Rigby, P. N., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 236-251 (1977).
15. Sherr, C. J., Fedele, L. A., Donner, L. & Turek, L. P. *J. Virol.* **32**, 860-875 (1979).
16. Barbacid, M., Lauver, A. V. & Devare, S. G. *J. Virol.* **33**, 196-207.
17. Ruscetti, S. K., Turek, L. P. & Sherr, C. J. *J. Virol.* **35**, 259-264 (1980).
18. Van de Ven, W. J. M., Reynolds, F. H. & Stephenson, J. R. *Virology* **101**, 185-197 (1980).
19. Barbacid, M., Beemon, K. & Devare, S. G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5158-5162 (1980).
20. Donner, L., Turek, L. P., Ruscetti, S. K., Fedele, L. A. & Sherr, C. J. *J. Virol.* **35**, 129-140 (1980).
21. Collett, M. S., Brugge, J. S. & Erikson, R. L. *Cell* **15**, 1363-1369 (1978).
22. Oppermann, H., Levinson, A. D., Varmus, H. E., Levintow, L. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1804-1808 (1979).
23. Collett, M. S., Erikson, E., Purchio, A. F., Brugge, J. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3159-3163 (1979).
24. Gardner, M. B. *et al. Nature* **226**, 807-809 (1970).
25. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).

Causes of more frequent deletions than insertions in mutations and protein evolution

Wilfried W. de Jong* & Lars Rydén†

* Department of Biochemistry, University of Nijmegen, 6525 EZ Nijmegen, The Netherlands

† Institute of Biochemistry, Biomedical Center, University of Uppsala, S-751 23 Uppsala, Sweden

Deletions and insertions of base pairs in DNA, jointly called gap events, are one of the major sources of evolutionary change at the molecular level. On the basis of very limited data, Fitch¹ has suggested that deletions might be expressed more often than insertions in proteins. We have examined the presently available homologous protein sequences, and observed that in the course of evolution, deletions of amino acids have indeed occurred about four times more frequently than insertions. An even higher preponderance of deletions over insertions (ten times or more) is found in recent spontaneous and induced mutants of genes or proteins. We propose here a mutational mechanism to explain this predominance of deletions.

Gap events are relatively rare in protein evolution, and mainly occur in the early stages of divergent evolution of duplicated genes². A study of gap events in the evolution of a protein requires a phylogenetic tree that allows us to differentiate between insertions and deletions. This means that the gap events must be located far enough from the root of the tree, and that the sequence homology must be sufficiently close to allow a reasonably unambiguous placement of gaps in the aligned sequences. In trees of distantly related proteins the large number of point mutations usually does not allow an unambiguous positioning of the gaps in the tree, whereas trees of closely related proteins have most of the gaps at the root. Therefore rather few protein families meet the above conditions for assignment of deletions or insertions.

We have, however, been able to trace nine families, comprising a total of 178 sequences, which are useful in this respect (Table 1). They contain a total of 37 gaps representing, on average, 2.2% of the total number of mutations. The figure

varies from less than 1% for trees of closely related sequences to 5% for more distantly related ones. As silent mutations are not included, the percent of gaps among the accepted changes must be even lower at the DNA level. Of the 37 gaps, 23 involve a single amino acid residue, 7 involve 2 residues and the remaining gaps involve 3 and 5 residues. The data show a fourfold excess of deletions over insertions. This preponderance of deletions in protein evolution suggests either that deletions are more readily accepted than insertions, or that there is a higher incidence of deletions than insertions at the DNA level. The former possibility is difficult to assess, but the latter explanation can be tested by examining gap events among mutant proteins and genes which are not constrained by selective pressure at the protein level.

Most informative in this respect are the abnormal human haemoglobins, of which a total of 204 different variants have been characterized³. Among these are 10 deletions, varying in length from one to five residues, and only a single insertion, of three residues (Table 2). Although numerous mutants of prokaryotic and eukaryotic genes and proteins are known, only very few have been characterized at the level of the nucleotide or amino acid sequence. However, large collections of mutants of the *lac I* gene and the λ -receptor propeptide gene of *Escherichia coli* have been analysed by nucleotide sequencing (Table 2). These studies show that >90% of the gap events are deletions, and that the gaps represent at least some 50% of the mutations when selection is absent.

We have also included (Table 2) the recently sequenced haemoglobin pseudogenes, which were probably functional genes up to a point where frameshift mutations and other changes prevented their expression, after which the acceptance of mutations in such pseudogenes was no longer constrained by selective forces at the protein level, and they were more free to accumulate deletions and insertions than coding genes. In the pseudogenes for which sequences have been published, 14 deletions and 7 insertions can be demonstrated with certainty. Comparison of other published sequences of non-coding DNA reveals that deletions, insertions and small duplications are indeed the most frequent mutational events in unconstrained sequences⁴, but does not allow any distinction between deletions and insertions.

The data show that the surplus of deletions and the preponderance of gaps among the mutations are inherent in the mutational process and that selection at the protein level reduces their proportion. Possible mechanisms for the creation of gaps include unequal cross-over between homologous genes^{1,5}, and,

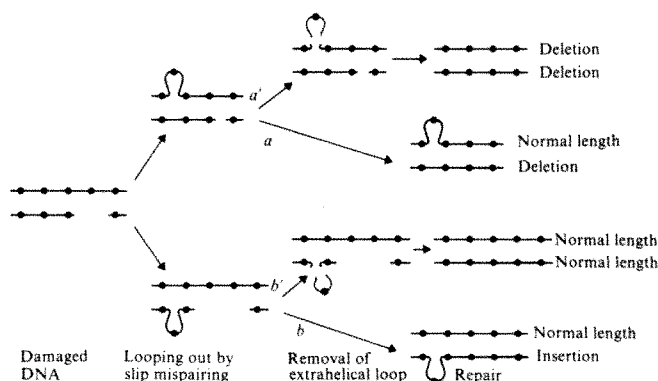


Fig. 1 Mutational model to explain the excess of deletions. Breakage of a DNA strand is followed by gap formation and looping-out of bases by slip mispairing; *a* and *b* represent the repair pathways resulting in deletion and addition of bases, respectively, according to the original Streisinger model⁹. If, as shown in *a'* and *b'*, the extrahelical loops are trimmed off by single-strand specific nuclease an excess of deletions follows. The removal of extrahelical bases can occur concomitantly with or after the repair process, but should take place before the replication process starts.

Table 1 Gap events in evolving proteins

Protein family	No. of sequences	No. of positions	No. of deletions	No. of insertions	No. of undecided gaps	Gaps as % of mutations	Ref.
Vertebrate myoglobins	31	153	1	0	0	0.5	15
Vertebrate pancreatic ribonucleases	43	123	2	0	2	1.2	16
Vertebrate α -crystallin A chains	43	173	2	0	0	2.0	17
Eukaryotic cytochrome c	6	101	2	0	2	2.4	
Eukaryotic and cyanobacterial plastocyanins	12	102	1	0	1	1.3	18
Vertebrate phospholipases	8	126	6	3	2	4.7	18
Mammalian lipotropins	4	93	0	1	0	3.3	18
Mammalian proinsulins	7	87	3	0	0	4.9	18
Vertebrate globins	24	146	7	2	0	2.5	18
Sum	178		24	6	7	2.2	

Phylogenetic trees, if not published, were constructed for each protein family¹⁴, and the number of substitutions, deletions, insertions and undecided gaps counted. Gaps were counted as a 21st amino acid. In the calculation of mutations the termini of the aligned polypeptide chains counted from the shortest one or a penultimate gap in the longer one were excluded. This was because gaps in this region might be due to base substitutions involving initiation or termination codons, and not to deletions or additions of bases, and furthermore, at penultimate positions the assignment of gaps or substitutions may be arbitrary. Several protein families that actually contain considerable numbers of gap events are not included because their cladistic relationships are difficult to assess (all viral and prokaryotic sequences; snake neurotoxins and cytotoxins; immunoglobulins) or their alignment too ambiguous (fibrinopeptides).

Table 2 Gap mutations in mutant genes or proteins subject to no selection or only partial selection

Gene or protein	No. of mutants	Deletions	Insertions	Gaps as % of total mutations	Type of selection	Ref.
Human haemoglobin variants	204	10	1	5	Functional or sublethal protein	3
<i>E. coli lac I</i> gene	140	20*	1†	46‡	No selection	10
<i>E. coli</i> λ -receptor protein gene	27	13		48	No selection	19
<i>Neurospora crassa</i> glutamate dehydrogenase, revertants of am6 mutant	14	14		100§		20
Haemoglobin pseudogenes :						
mouse α -3		8	6	18	Possibly no selection after expression ceased	21
human $\psi\alpha$ 1		3		3		22
rabbit $\psi\beta$ 2		2		3		23
goat β^x		1	1	4		24

* One of these mutations appeared 18 times in the collection, but is counted only once (hotspot mutation).

† This mutation appeared 78 times in the collection but is counted only once (same hotspot as in *). It reverted with a frequency of 10^{-5} .

‡ This is a low estimate because not all mutants were characterized. If the multiple hotspot mutations are included it increases to >80%.

§ All 14 revertants analysed were found to be deletions, although the am6 mutant itself is thought to be a one-base insertion.

|| Only the positions corresponding to the expressed portions of the functional genes are considered.

in some cases, changes at intron-exon junctions⁶⁻⁸. Most gaps, however, can probably be adequately explained by the Streisinger model⁹ in which DNA strand breakage is followed by 'slipped mispairing' and consolidation of the mispaired configuration by DNA repair. Such 'slippage' would be favoured by tandemly repeated nucleotide sequences; insertions and deletions frequently (but not always) occur in repeated sequences^{5,9-11}. In agreement with this model, it has been found that most gap mutations produced by UV light, arise during excision repair⁵.

The Streisinger model explains deletions equally well as insertions, and does not *a priori* lead one to expect a preponderance of either. However, a simple extension of this model can explain an excess of deletions. In the Streisinger model the slipped mispairing after strand breakage results in the looping-out of a number of bases. Such transient extrahelical loops should be vulnerable to single-strand specific nuclease action¹², and might be trimmed off (and the resulting broken strand ligated) before replication occurs. Such a process would decrease the frequency of insertions and favour the occurrence of deletions (Fig. 1). Similarly, the recently proposed mechanism for the generation of deletions by slipped mispairing during replication¹¹ could contribute to the scarcity of insertions.

A higher incidence of deletions than insertions during protein evolution should lead to a gradual decrease in the lengths of

polypeptide chains—this seems to be true for the insulin C peptide¹³, and the ancestral globin chain was probably longer than the contemporary vertebrate haemoglobin and myoglobin chains (excluding terminal extensions)². However, the shortening effect of the deletion events is not dramatic: they are very rarely accepted in proteins with stabilized functions, compared with the frequency of amino acid substitutions.

Received 8 December 1980; accepted 26 January 1981.

- Fitch, W. M. *A. Rev. Genet.* **7**, 343-380 (1973).
- Dayhoff, M. O. (ed). *Atlas of Protein Sequence and Structure* Vol. 5 (National Biomedical Research Foundation, Washington DC, 1972).
- Hemoglobin 3, 104-115, 376-395, 482-513 (1979).
- Konkel, D. A., Maizel, J. V. & Leder, P. *Cell* **18**, 865-873 (1979).
- Hanawalt, P. C., Cooper, P. K., Ganesan, A. K. & Smith, C. A. *A. Rev. Biochem.* **48**, 783-836 (1979).
- Wallis, M. *Nature* **284**, 512 (1980).
- De Jong, W. W., Cohen, L. H., Leunissen, J. A. M. & Zweers, A. *Biochem. biophys. Res. Commun.* **96**, 648-655 (1980).
- Stein, J. P., Catterall, J. F., Kristo, P., Means, A. R. & O'Malley, B. W. *Cell* **21**, 681-687 (1980).
- Streisinger, G. *et al. Cold Spring Harb. Symp. quant. Biol.* **31**, 77-84 (1966).
- Farabaugh, P. J., Schmeissner, U., Hofer, M. & Miller, J. H. *J. molec. Biol.* **126**, 847-863 (1978).
- Elstratidis, A. *et al. Cell* **21**, 653-668 (1980).
- Dodgson, J. B. & Wells, R. D. *Biochemistry* **16**, 2374-2379 (1977).
- Tager, H. S. & Steiner, D. F. *J. biol. Chem.* **247**, 7936-7940 (1972).
- Fitch, W. M. & Margoliash, E. *Science* **155**, 279-284 (1967).

15. Romero-Herrera, A. E., Lehmann, H., Joysey, K. A. & Friday, A. E. *Phil. Trans. R. Soc. B* **283**, 61–163 (1978).
16. Beintema, J. J. & Lenstra, J. A. in *Macromolecular Sequences in Systematics and Evolutionary Biology* (ed. Goodman, M.) (Plenum, New York, in the press).
17. De Jong, W. W., Zweers, A. & Goodman, M. *Colloq. Prot. Biol. Fluids* Vol. 28, 161–164 (Pergamon, Oxford, 1980).
18. Dayhoff, M. O. (ed.) *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (National Biomedical Research Foundation, Washington DC, 1978).
19. Emr, S. D., Hedgpeth, J., Clément, J.-M., Silhavy, T. J. & Hofnung, M. *Nature* **285**, 82–85 (1980).
20. Sidding, M. A. M., Kinsey, J. A., Fincham, J. R. S. & Keighren, M. *J. molec. Biol.* **137**, 125–135 (1980).
21. Nishioka, Y., Leder, A. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2806–2809 (1980).
22. Proudfoot, N. J. & Maniatis, T. *Cell* **21**, 537–544 (1980).
23. Lacy, E. & Maniatis, T. *Cell* **21**, 545–553 (1980).
24. Cleary, M. L., Haynes, J. R., Schon, E. A. & Lingrel, J. B. *Nucleic Acids Res.* **8**, 4791–4802 (1980).

Substituted benzimidazoles inhibit gastric acid secretion by blocking ($H^+ + K^+$)ATPase

Erik Fellenius*, Thomas Berglindh†, George Sachs‡, Lars Olbe§, Berit Elander*, Sven-Erik Sjöstrand* & Björn Wallmark*

* Research Laboratories, AB Hässle, S-43183 Mölndal, Sweden

† Department of Physiology and Medical Biophysics, Biomedical Centre, Uppsala University, S-75105 Uppsala, Sweden

‡ Laboratory of Membrane Biology, University of Alabama in Birmingham, Birmingham, Alabama 35294, USA

§ Sahlgrenska Hospital, Surgical Clinic II, S-41345 Gothenburg, Sweden

Studies both *in vivo*^{1,2} and *in vitro*^{1–5} have shown that substituted benzimidazoles inhibit the stimulation of acid secretion produced by dibutyryl cyclic AMP and histamine. Furthermore, the results differ from those produced by H_2 antagonists and anticholinergic agents in that the inhibition is not competitive, and the site of action is intracellular and peripheral to that of dibutyryl cyclic AMP. To investigate the biochemical mechanism of action of substituted benzimidazoles, one such compound, H 149/94 (2-[2-(3-methylpyridyl-methyl)-sulphonyl]-5-methoxycarbonyl-6-methylbenzimidazol), has been tested either directly on an ($H^+ + K^+$)ATPase isolated from pig and human gastric mucosa or on the function of this enzyme in gastric glands isolated from rabbit and human gastric mucosa. ($H^+ + K^+$)ATPase^{6,7}, which has only been found at the secretory surface of the parietal cell⁸, catalyses a one-to-one exchange of protons and potassium ions^{9–11}. It is possibly the proton pump within the gastric mucosa, and may thus be the terminal or one of the terminal steps of the acid secretory process^{12,13}. We show here that H 149/94 inhibits ($H^+ + K^+$)ATPase, which may explain its inhibitory action on acid secretion *in vitro* and *in vivo*. Because of the unique distribution and properties of the ($H^+ + K^+$)ATPase, the inhibitory action of H 149/94 on this enzyme may be a highly selective clinical means of suppressing the acid secretory process.

Isolated gastric glands have been shown to respond to a number of different secretagogues and inhibitors^{5,14,15}. Secretagogues such as dibutyryl cyclic AMP and K^+ stimulate acid secretion intracellularly, presumably at a site close to the ($H^+ + K^+$)ATPase^{16–18}. In contrast, histamine and acetylcholine have a site of action further removed from the enzyme^{14–16,19}. This may be either extracellular or at an earlier stage in the sequence of events leading to acid secretion, possibly reflecting the physiological role of such agonists in the control of acid secretion. Thus, isolated gastric glands, used in conjunction with specific cholinergic and histaminergic antagonists, provide a useful experimental model for determining the site of action of drugs influencing acid secretion, independently of hormonal and

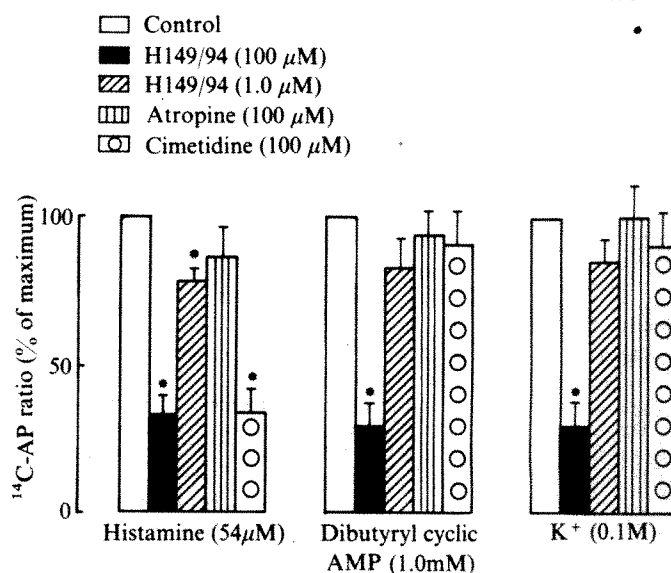


Fig. 1 The inhibitory action of H 149/94, atropine and cimetidine on ^{14}C -AP accumulation induced by different agonists in isolated gastric glands from rabbit mucosa. The preparation and incubation of the isolated glands was essentially as described elsewhere^{14,15} except that the incubation volume was reduced from 3 to 0.2 ml and the amount of glands also reduced proportionally. The glands were suspended in a physiological buffer solution containing rabbit albumin (2 mg ml⁻¹), glucose (10 mM) and ^{14}C -AP (0.25 μCi and 8 μmol per incubation), and incubated in a shaking bath for 90 min at 37°C. After the incubation the glands were spun down at 2,000g for 2 min. The supernatant was removed and the wet and dry weights of the pellet determined. The dried pellets were solubilized in 1 ml Soluene. Dimilume (10 ml) was added and the radioactivity then counted. H 149/94 was dissolved in methanol. All incubations contained 1% methanol, which had no influence on the ^{14}C -AP ratio. All other added substances were dissolved in the physiological buffer solution. In the experiments with high [K^+] (0.1 M), the glands were washed three times with an incubation medium in which KCl (0.1 M) had been substituted for NaCl (0.1 M). The respective control values without antagonist have been set to 100%. The results are mean $\pm$ s.e. of 6–8 observations. A statistically significant difference ($P < 0.05$ Student's *t*-test) from control is denoted by an asterisk.

nervous influence. Although the rate of acid secretion cannot be measured quantitatively with isolated glands, the acid secretory response can be monitored semiquantitatively by measuring the uptake of a radiolabelled weak base, ^{14}C -aminopyrine (^{14}C -AP), which accumulates in the glands in proportion to the pH differences between the intraglandular acid compartments and the surrounding medium^{14,15}.

The effects of H 149/94 (10^{-4} , 10^{-6} M), cimetidine (10^{-4} M) and atropine (10^{-4} M) on ^{14}C -AP accumulation in rabbit gastric glands stimulated by histamine, dibutyryl cyclic AMP and K^+ are shown in Fig. 1. Atropine had no inhibitory effect in the presence of any agonist. Cimetidine inhibited the response to histamine (54 μM) but had no effect in the presence of dibutyryl cyclic AMP (1 mM) or K^+ (0.1 M). H 149/94 produced dose-dependent inhibition ($ED_{50} \sim 10^{-5}$ M) of ^{14}C -AP accumulation irrespective of the agonist used. The presence of potassium is an absolute requirement for acid secretion, presumably reflecting its involvement in the exchange process catalysed by ($H^+ + K^+$)ATPase^{17,18}. The marked inhibition of K^+ -induced ^{14}C -AP accumulation produced by H 149/94 therefore strongly supports the view that this compound interferes with acid secretion at or close to the enzyme, the probable terminal step of the secretory process. The lack of effect with cimetidine and atropine suggests a more distant site of action of these compounds.

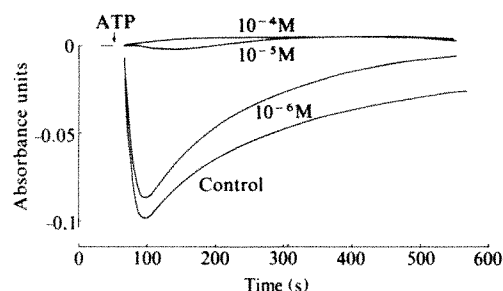


Fig. 2 Effect of H 149/94 on proton transport catalysed by the ($H^+ + K^+$)ATPase. Gastric membranes were prepared from pig stomach by differential and zonal density gradient centrifugation²⁰. The vesicle fraction ($\sim 40 \mu\text{g ml}^{-1}$) was equilibrated at 4°C overnight in the following medium: 2 mM PIPES-Tris, pH 6.7, 150 mM KCl, 2 mM MgCl_2 with or without H 149/94 at concentrations of 10^{-4} , 10^{-5} , 10^{-6} M as indicated. H 149/94 was dissolved in methanol. All incubations contained 1% methanol, which had no influence on the assay. At the start of the experiment $20 \mu\text{M}$ acridine orange was added to the sample and the reaction started by the addition of 0.6 mM ATP (disodium salt). The acridine orange absorbance change was monitored in an Aminco DW-2 spectrophotometer in the dual-beam mode set at 490 and 530 nm.

To explore the possibility that H 149/94 interacts directly with ($H^+ + K^+$)ATPase, the enzyme was isolated from pig stomach by differential and zonal gradient centrifugation²⁰. The vesicular fraction sedimenting between the 0.25 M sucrose–8% Ficoll interface was either used directly for proton transport studies, or further purified by free-flow electrophoresis and subsequent lyophilization for studies with an ATPase assay. Fresh vesicles are relatively impermeable to cations such as protons and K^+ and readily allow measurements of proton transport⁹, which can be quantified using the dye acridine orange. The control curve in Fig. 2 shows the acridine orange absorbance, when ATP is added to the K^+ -preloaded vesicles in

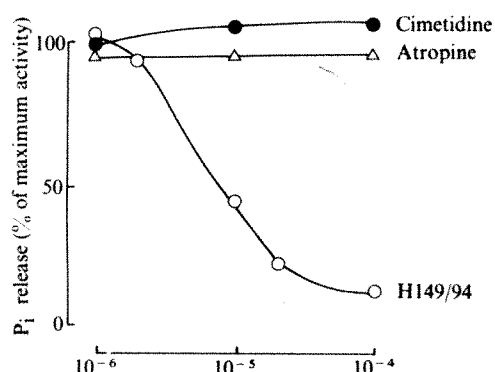


Fig. 3 The effects of H 149/94, atropine and cimetidine on isolated gastric ($H^+ + K^+$)ATPase from pig mucosa. The ($H^+ + K^+$)ATPase was obtained by differential and zonal density discontinuous gradient centrifugation followed by free-flow electrophoresis²⁰. ($H^+ + K^+$)ATPase activity was quantified by the K^+ -stimulated release of phosphate²². The incubation concentrations were Tris-HCl 40 mM, pH 7.4, MgCl_2 2 mM, ATP 2 mM, with or without KCl 10 mM. The enzyme ($\sim 10 \mu\text{g}$ per incubation) was preincubated with the drugs for 30 min and the time of incubation was 15 min at 37°C . H 149/94 was dissolved in methanol. All incubations contained 1% methanol, which had no influence on ($H^+ + K^+$)ATPase activity.

the absence of inhibitor. The interpretation of the control curve has been described elsewhere^{11,21} but briefly, on addition of ATP there is a rapid loss of acridine orange absorbance corresponding to transport of H^+ into the vesicle. This is followed by an increase in the absorbance produced by an outward H^+ leak, due to depletion of intravesicular K^+ . Thus, the H^+ transport signal is both ATP and K^+ dependent. In this experiment (Fig. 2) the vesicle preparation was incubated with H 149/94 at concentrations of 10^{-6} , 10^{-5} and 10^{-4} M. There was progressive inhibition of the transport signal with increasing H 149/94 concentrations—that is, proton transport was inhibited.

To test whether this inhibition of proton transport was a consequence of decreased ($H^+ + K^+$)ATPase pump activity, K^+ -ATPase activities were measured at different H 149/94 concentrations. Figure 3 shows that H 149/94 produces dose-dependent inhibition of ($H^+ + K^+$)ATPase activity in contrast to cimetidine and atropine. We therefore conclude that the inhibition of proton transport by H 149/94 (Fig. 2) corresponds to the inhibition of ($H^+ + K^+$)ATPase activity seen in Fig. 3.

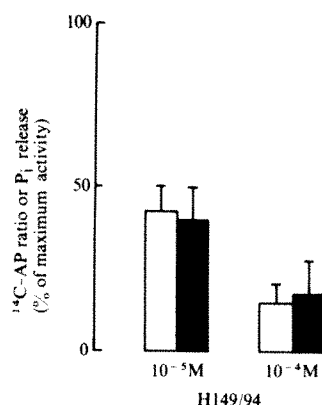


Fig. 4 The inhibitory action of H 149/94 on dibutyl cyclic AMP (10^{-3} M)-induced ^{14}C -AP accumulation in human isolated gastric glands ($\square$) and P_i release catalysed by isolated human ($H^+ + K^+$)ATPase ($\blacksquare$). The mucosa was obtained from surgical biopsies taken during gastrectomy for either recurrent gastric ulcer or recurrent ulceration after previous vagotomy. The isolated glands⁵ were prepared as for the isolated gastric glands from rabbit (Fig. 1). ($H^+ + K^+$)ATPase was prepared by differential centrifugation and further purified on a discontinuous step gradient⁵. The ATPase assay was performed as for pig ($H^+ + K^+$)ATPase (Fig. 3) except that the release of radiolabelled phosphate from $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ was measured²³. The results are mean $\pm$ s.e. of four observations.

H 149/94 inhibits gastric acid secretion in man (Olbe *et al.*, unpublished observation). As the therapeutic value of the compound may be its antisecretory activity in peptic ulcer patients, a series of experiments similar to those described above was performed on isolated gastric glands and ($H^+ + K^+$)ATPase obtained from gastric mucosal biopsy specimens from such patients. The results are shown in Fig. 4. H 149/94, at concentrations comparable with those found effective in the animal preparations, produced dose-dependent inhibition of both dibutyl cyclic AMP-induced ^{14}C -AP accumulation in gastric glands, and P_i release from vesicles.

Thus, these results suggest that the substituted benzimidazole, H 149/94, is an inhibitor of the ($H^+ + K^+$)ATPase and provide a mechanism for the inhibitory effects of H 149/94 on acid secretion which have been demonstrated *in vitro* and *in vivo*¹⁻⁵. Because of the unique distribution of the ($H^+ + K^+$)ATPase, the inhibitory effects of H 149/94 may offer a highly selective means of suppressing secretion for therapeutic purposes.

Received 25 November 1980; accepted 6 February 1981.

1. Fellenius, E. *et al.* in *Hydrogen Ion Transport in Epithelia* (eds Schulz, I., Sachs, G., Forte, J. & Ullrich, K. J.) 193–202 (Elsevier, Amsterdam, 1980).
2. Olbe, L., Sjöstrand, S.-E. & Fellenius, E. in *Gastrins and the Vagus* (eds Rehfeld, J. F. & Andrup, E.) 245–250 (Academic, New York, 1979).
3. Sjöstrand, S.-E., Ryberg, B. & Olbe, L. *Acta physiol. scand. spec. Suppl.*, 181–185 (1978).
4. Olbe, L. *et al. Scand. J. Gastroenterol. Suppl.* **55**, 131–133 (1979).
5. Fellenius, E. *et al.* in *Hormone Receptors in Digestion and Nutrition* (eds Rosselin, G., Fromageot, P. & Bonfils, S.) 355–360 (Elsevier, Amsterdam, 1979).
6. Lee, J., Simpson, G. & Scholes, P. *Biochem. biophys. Res. Commun.* **60**, 825–832 (1974).
7. Ray, T. K. & Forte, J. G. *Biochim. biophys. Acta* **443**, 451–467 (1976).
8. Saccomani, G. *et al. J. Cell Biol.* **83**, 271–283 (1979).
9. Shackman, R., Schwartz, A., Saccomani, G. & Sachs, G. *J. Membrane Biol.* **32**, 361–381 (1977).
10. Sachs, G., Chang, H. M., Rabon, E., Shackman, R. & Saccomani, G. *J. biol. Chem.* **251**, 7690–7698 (1976).
11. Chang, H., Saccomani, G., Rabon, E., Shackman, R. & Sachs, G. *Biochim. biophys. Acta* **464**, 313–327 (1977).
12. Sachs, G. *et al. Gastroenterology* **71**, 931–931 (1977).
13. Forte, J. G. & Lee, H. C. *Gastroenterology* **73**, 921–927 (1977).
14. Berglindh, T. & Öbrink, K.-J. *Acta physiol. scand.* **96**, 150–159 (1976).
15. Berglindh, T. *Biochim. biophys. Acta* **464**, 217–233 (1977).
16. Chew, C. S., Hersey, S. J., Sachs, G. & Berglindh, T. *Am. J. Physiol.* **238**, 312–320 (1980).
17. Berglindh, T., Helander, H. F. & Sachs, G. *Scand. J. Gastroenterol. Suppl.* **55**, 7–15 (1979).
18. Berglindh, T., Dibona, D. R., Ito, S. & Sachs, G. *Am. J. Physiol.* **238**, G 165–G 176 (1980).
19. Soll, A. *Am. J. Physiol.* **238**, G 366–G 375 (1980).
20. Saccomani, G., Stewart, H. B., Shaw, D., Lewin, M. & Sachs, G. *Biochim. biophys. Acta* **465**, 311–330 (1977).
21. Rabon, E., Chang, H. & Sachs, G. *Biochemistry* **17**, 3345–3353 (1978).
22. Fiske, C. H. & Subbarow, Y. *J. Biol. Chem.* **66**, 375–400 (1925).
23. Mårdh, S. *Biochim. biophys. Acta* **391**, 448–463 (1975).

Location of hexane in lipid bilayers determined by neutron diffraction

Stephen H. White*†, Glen I. King*
& James E. Cain ‡§

*Department of Physiology and Biophysics, University of California, Irvine, California 92717, USA and the Biology Department, Brookhaven National Laboratory.

‡Biology Department, Brookhaven National Laboratory, Upton, New York 11973, USA

Knowledge of the interactions of alkanes and other small hydrophobic molecules dissolved in lipid bilayers is important for understanding lipid–protein interactions in membranes¹, the microscopic properties of solutions^{2,3} and the mechanism of anaesthesia^{4–6}. An essential element for describing these interactions is the distribution of the molecules across the thickness of the bilayer. Studies of black lipid films^{1,2,7–9} strongly suggest that dissolved alkanes are located primarily in the centre of the bilayer. Recent X-ray and neutron diffraction and thermodynamic measurements on lipid dispersions and oriented lipid multilayers are consistent with this view^{10–12}. We present here the first direct evidence for this hypothesis obtained from neutron diffraction studies of oriented dioleoyl lecithin (DOL) multilayers containing deuterated hexane introduced using the vapour phase. The hexane is found mainly in a zone 10 Å wide in the centre of the bilayer. We have also estimated the amount of hexane in the bilayers and the free energy of transfer from pure hexane to bilayer. Our numbers are considerably different from those reported by Simon *et al.*¹³ for the interaction of hexane with DOL liposomes. The differences may be due to difference in water activity and bilayer curvature¹⁴.

The neutron scattering length of the hydrogen nucleus is -0.38×10^{-12} cm whereas for the deuterium nucleus it is $+0.65 \times 10^{-12}$ cm (ref. 15). Consequently, the introduction of deuterated alkane into a bilayer containing non-deuterated acyl chains produces large changes in the neutron scattering density of the hydrophobic core of the bilayer. The location of the alkane can be established in the following way. Diffraction measurements are performed on lipid bilayers containing hydrogenated alkane

(H-alkane) and on bilayers containing deuterated alkane (D-alkane) at the same concentration. The resulting structure factors are used to construct neutron scattering density profiles $\rho_H(x)$ and $\rho_D(x)$ of the bilayers containing, respectively, H-alkane and D-alkane. Subtraction of the two profiles yields the difference scattering density profile $\Delta\rho(x) = \rho_D(x) - \rho_H(x)$. The contribution made to each profile by the water, polar groups and acyl chains are the same and cancel out. Therefore, $\Delta\rho(x)$ represents the distribution of the deuterated alkane across the thickness of the bilayer. This general method¹⁶ has been used to locate water^{17,18}, cholesterol¹⁸ and specific atoms on phosphatidylcholine molecules^{19–21} in lipid bilayers.

Here we have measured the distribution of *n*-hexane in DOL bilayers at relatively low hexane concentrations. Hexane was chosen for this initial study of alkane–bilayer interactions because (1) it has a high vapour pressure and can be introduced into the bilayer via the vapour phase in equilibrium conditions; (2) the vapour pressure can be controlled with hexane/hexadecane mixtures whose thermodynamic properties are well

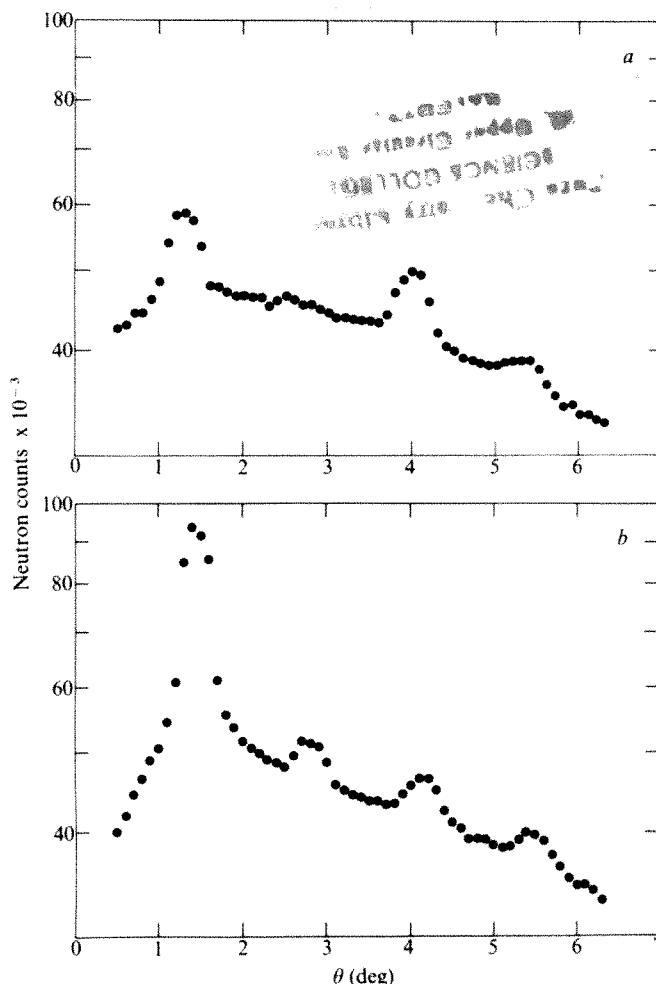


Fig. 1 Diffraction patterns collected at the High Flux Beam Reactor at Brookhaven National Laboratory of the lamellar reflections of DOL multilayers in the presence of deuterated hexane (a) and hydrogenated hexane (b). Both patterns show four orders of diffraction ($h=4$) as recorded by a two-dimensional detector in fixed position as the sample was rotated in 0.1° steps through the Bragg angle with respect to the incident beam. The diffraction patterns were the typical lamellar type with a mosaic spread of 0.1 – 0.4° . The typical d -spacing was 48 Å. The sample was oriented on a quartz slide sealed in an aluminium chamber (temperature 22.5°C). RH was maintained at 66% with a saturated sodium nitrite solution ($100\% \text{H}_2\text{O}$, $0\% \text{D}_2\text{O}$). The vapour pressure (P) of hexane was fixed at 40 mm Hg with a hexane/hexadecane mixture. $P = X_H^m P_0$ where X_H^m is the mole fraction of hexane in the mixture (0.3 in this case) and P_0 is the vapour pressure of pure hexane (134 mm Hg, 22.5°C). The sample came into equilibrium with the hexane vapour within 2 h as shown by unchanging diffraction patterns.

† To whom correspondence should be addressed (at the University of California).

§ Present address: Solid Photography, Inc., Melville, New York 11746, USA.

known²²; (3) thermodynamic information is available on the hexane-DOL interaction in liposomes¹³ and similar information can be obtained for the oriented multibilayer system used here; and (4) hexane is an anaesthetic agent²³.

DOL multilayers were deposited on a quartz slide by evaporation from a 1:1 chloroform/methanol solution and mounted in the neutron beam in a sealed aluminium container. The DOL had been purified by thick-layer chromatography and great care was taken to insure that all the solvent used in depositing the sample had been removed. The experiments were performed at 22.5 °C and 66% relative humidity (RH) maintained by a saturated aqueous solution of sodium nitrite. The vapour pressure, P , of the hexane was controlled by hexane/hexadecane mixtures which behave nearly ideally so that $P = X_H^m P_0$ where P_0 is the vapour pressure of pure hexane (134 mm Hg at 22.5 °C) and X_H^m is the mole fraction of hexane in the mixture. X_H^m ranged from 0 to 0.3 for these experiments. The hexane activity (X_H^m) required for narcosis in mice is 0.13 at 37 °C (ref. 23).

Typical uncorrected $I(\theta)$ curves collected from DOL multilayers in the presence of hexane are shown in Fig. 1. The upper curve (a) is for D-hexane and the lower curve (b) for H-hexane. In both cases, $X_H^m = 0.3$. The two $I(\theta)$ curves are clearly different due to the difference in scattering density of H- and D-hexane. As shown in Fig. 2, the structure factors derived from such data were accurately a linear function of the percentage of D-hexane in the hexane phase. The phases for the first four orders for $X_H^m = 0.3$ are -, -, +, - and are the same as for

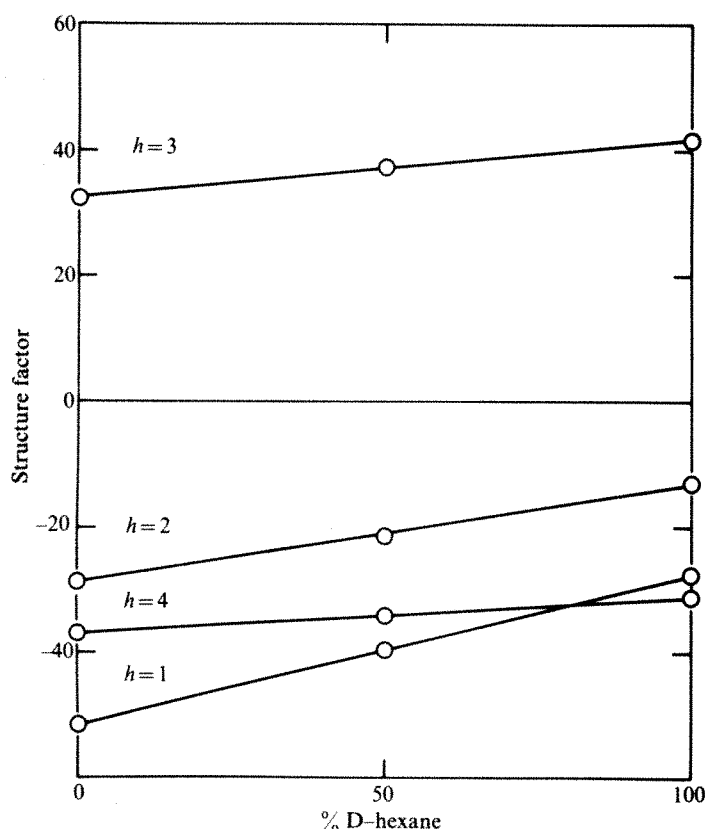


Fig. 2 Structure factors $F(h)$ for DOL multilayers in the presence of hexane vapour as a function of the per cent D-hexane in the vapour phase (temperature 22.5 °C, RH = 66% [100% H₂O, 0% D₂O], $X_H^m = 0.3$ [$P = 40$ mm Hg]). $F(h) = s(h) (I_h)^{1/2}$ where I_h is the intensity corrected for background and sample-beam intersection and $s(h)$ is the correct phase assignment (+1 or -1). The Lorentz correction has also been applied. The $F(h)$ values are linear, justifying the use of the difference Patterson deconvolution method^{16,25} to establish the phase assignment for each value of F . There were no changes in phase as X_H^m varied from 0 to 0.3. The figure shows the phases to be the same for D-hexane and H-hexane in this range. The phase assignments for the first four orders are -, -, +, -.

$X_H^m = 0$. These phases agree with assignments by others for dipalmitoyl lecithin¹⁷, dipalmitoyl lecithin-cholesterol¹⁶ and egg lecithin-cholesterol¹⁸.

Figure 3a shows the neutron scattering density profiles $\rho_D(x)$ and $\rho_H(x)$ for DOL bilayers containing, respectively, D-hexane (curve D) and H-hexane (curve H) at $X_H^m = 0.3$. The H-hexane profile has the typical appearance expected of bilayers and is essentially identical to profiles of hexane-free bilayers (not shown). This result is consistent with the findings of Franks and Lieb¹¹ on lecithin/cholesterol bilayers containing cyclopropane—the alkanes at moderate concentrations have no significant effect on the gross structure of the bilayer. However, the presence of alkane in the bilayer is clearly shown by the D-hexane profile. Note that the methyl trough in the centre of the bilayer has been 'filled in' due to the presence of the D-hexane in the bilayer centre. This is shown more dramatically in Fig. 4 in which difference profiles $\rho_D(x) - \rho_H(x)$ are plotted for $X_H^m = 0.05, 0.10, 0.20$ and 0.30. These profiles unambiguously demonstrate that the hexane is largely confined to the centre of the bilayer and that the amount of hexane in the centre depends directly on hexane vapour pressure.

A surprising finding is that the d -spacing of the samples does not depend, within the instrumental sensitivity (± 0.4 Å), on the hexane vapour pressure. Consistent with this finding, Franks and Lieb¹¹ reported that halothane at high concentrations has no effect on the thickness of lecithin/cholesterol bilayers. To examine further this finding and to estimate the amount of hexane in the bilayer, we have constructed absolute neutron scattering density 'strip' models for the bilayers for the different experimental conditions to ~ 10 Å resolution (Fig. 3b). Acceptable models predicted structure factors that agreed with the experimental ones to within 1%. We made several interesting observations using this approach. First, the hydrocarbon thickness of the bilayer is independent of X_H^m and apparently does not vary by more than 0.1 Å. Second, the hexane is confined largely to the centre of the bilayer in a zone ~ 10 Å thick, although at the highest X_H^m a very small amount of hexane begins to appear in the hydrophobic regions near the polar groups (Fig. 3a). Third, the scattering density of the central zone is, to within experimental error, a linear function of X_H^m . This last result means that we were performing the experiments in the Henry's law region of the bilayer's hexane solubility curve.

We have estimated the amount of D-hexane in the bilayers from the strip models. The mole fraction X_H^b of hexane in bilayer (relative to the number of acyl chains) is 0.105 for $X_H^m = 0.30$, corresponding to about 12 hexanes per 100 acyl chains. If there were no change in A_0 (the area per phospholipid at the lipid/water interface), this much hexane should have caused the hydrophobic core of the bilayer to thicken by about 1.3 Å. Our absolute neutron scattering density models, which are sensitive to changes in hydrocarbon thickness as small as 0.1 Å, show no thickening and we conclude that A_0 must increase from ~ 70 Å² to 74 Å² as X_H^m increases from 0 to 0.3. Alkanes cause large increases in the thickness of black lipid membranes^{1,7} without apparently affecting A_0 (ref. 7). We do not know why we observe no change in thickness but we note that, unlike the usual situation for black films, the activities of hexane and water were considerably less than 1 in our experiments. White²⁴ has measured the effect on bilayer thickness of variations in the mole fraction (X_D^b) of decane in the torus surrounding black films formed from glyceryl monooleate. Increasing X_D^b from 0 to 0.3 causes the thickness to increase by 1.2 Å, which suggests that water activity may be important in determining the response of A_0 to alkanes dissolved in the bilayer because the black film is formed at a water activity of almost 1.

Because we can estimate the amount of hexane in the DOL bilayers as a function of X_H^m , we can also estimate the free energy of transfer ($\mu_D^b - \mu_m^b$) of hexane from the pure liquid to the bilayer. Consistent with being in the Henry's law solubility region, we find that the ratio X_H^b/X_H^m is independent of X_H^m and is 2.86 ± 0.16 . This ratio is, of course, the activity coefficient of

the hexane in the bilayer referred to pure hexane. The hexane/hexadecane mixture is nearly ideal, so that $\mu_b^0 - \mu_m^0 = RT \ln (X_H^m/X_H^b)$, which yields 600 cal mol^{-1} . If one refers X_H^b to the number of phospholipids rather than the number of acyl chains, the free energy of transfer is 270 cal mol^{-1} , which is considerably smaller than the value of $1,560 \text{ cal mol}^{-1}$ reported by Simon *et al.*¹³ for DOL liposomes in water. Our results also differ from those of Simon *et al.*¹³ in that we estimate from their data that for $X_H^m = 0.30$, the mole fraction of hexane in the liposome bilayer is ~ 0.01 referred to the number of acyl chains; this is one-tenth the value obtained by us.

The large differences between our results and those of Simon *et al.*¹³ may exist for several reasons. First, our estimate of X_H^m may be in error because of the problems inherent in constructing absolute neutron scattering density models. A factor of 10 error

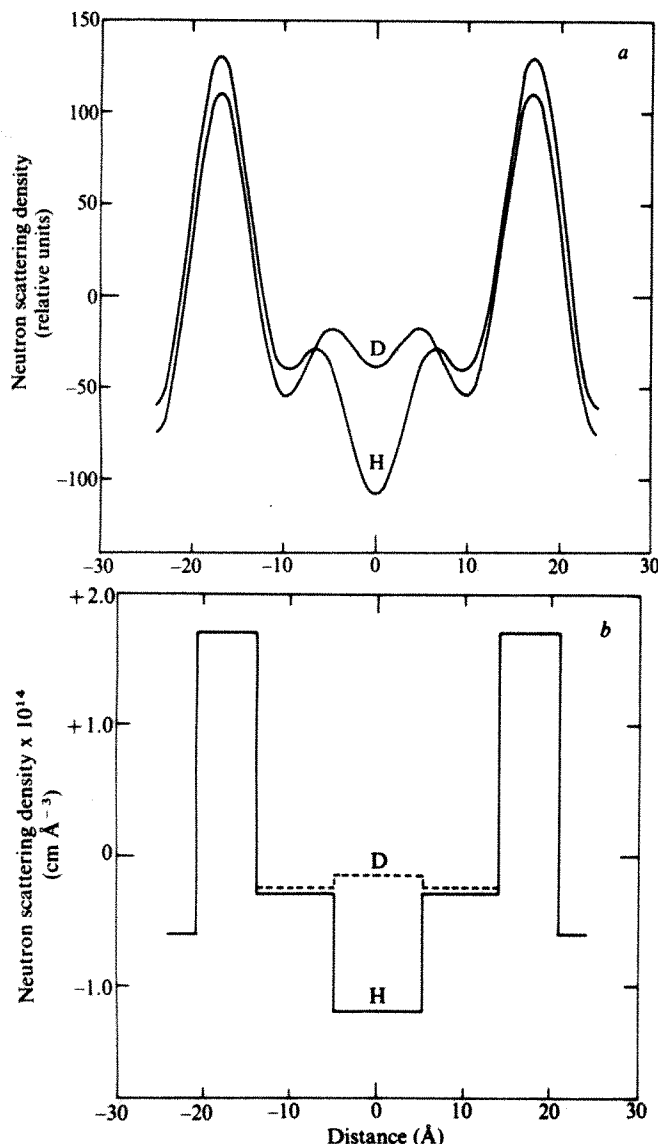


Fig. 3 *a*, 12 Å-resolution neutron scattering density profiles (relative scale) for oriented DOL multilayers containing D-hexane (D) and H-hexane (H) (temperature 22.5 °C, RH = 66% [100% H₂O, 0% D₂O], $X_H^m = 0.3$). The presence of the D-hexane is clearly revealed by the increased scattering density in the centre of the profile for the D-hexane case (D). *b*, Absolute neutron scattering density strip models whose 12 Å Fourier reconstructions agree with the syntheses in *a* with an *R*-value of <1%. The parameters of the model widths and density levels were varied to give the values shown. Similar models were constructed for the other X_H^m values, leading to the result that the central 27.4 Å hydrocarbon thickness apparently does not vary by more than 0.1 Å, assuming a constant *d*-spacing. However, as we were unable to determine the *d*-spacing with a precision of better than 0.4 Å, changes of this magnitude in the hydrocarbon thickness cannot be ruled out.

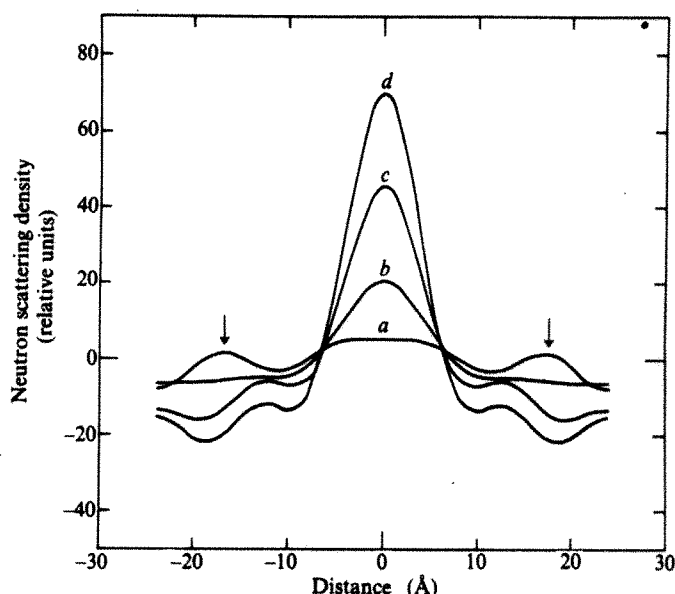


Fig. 4 Difference neutron scattering density profiles for oriented DOL multilayers containing different amounts of hexane (temperature 22.5 °C, RH = 66%). Each curve is the result of subtracting the profile of the bilayer containing H-hexane from the profile of the bilayer containing D-hexane at the same X_H^m . $X_H^m = 0.05$ (a), 0.10 (b), 0.20 (c) and 0.30 (d). These difference profiles show that the hexane is located in the centre of the bilayer and that the amount of hexane in the bilayer depends directly on X_H^m . The negative difference densities in the regions marked by the arrows are termination artefacts.

seems unlikely, however. Second, our oriented DOL multilayers exist at a water activity of 0.66 whereas the liposomes are formed in excess water. Third, the multilayered liposomes are curved whereas our oriented samples are planar. Gruen and Haydon¹⁴ have proposed that the curvature of liposomes causes a large average decrease in the solubility of alkanes in liposome bilayers. They estimate that the alkane absorption per mole of lipid in liposomes may be as small as 17% of the planar value. This is in reasonable agreement with our result taking the assumptions of their model into account.

This research was carried out at Brookhaven National Laboratory under the auspices of the Department of Energy with the additional support of the NSF. We acknowledge the advice and encouragement of Dr Benno P. Schoenborn and the help of his staff. The research was supported by grants from the NSF and NIH. S.H.W. was supported by a Research Career Development Award from the NIH.

Received 18 August; accepted 24 December 1980.

- White, S. H. *Ann. N.Y. Acad. Sci.* **303**, 243–265 (1977).
- White, S. H. *Science* **207**, 1075–1077 (1980).
- Simon, S. A., Stone, W. L. & Bennett, P. B. *Biochim. biophys. Acta* **550**, 38–47 (1979).
- Franks, N. P. & Lieb, W. R. *Nature* **274**, 339–342 (1978).
- Miller, K. W. & Pang, K.-Y. Y. *Nature* **263**, 253–255 (1976).
- Haydon, D. A., Hendry, B. M., Levinson, S. R. & Requena, J. *Nature* **268**, 356–358 (1977).
- Andrews, D. M., Manev, E. D. & Haydon, D. A. *Spec. Discuss. Faraday Soc.* **1**, 46–56 (1970).
- White, S. H. *Nature* **262**, 421–422 (1976).
- Brooks, D. E., Levine, Y. K., Requena, J. & Haydon, D. A. *Proc. R. Soc. A* **347**, 179–194 (1975).
- McIntosh, T. J., Simon, S. A. & MacDonald, R. C. *Biochim. biophys. Acta* **597**, 445–463 (1980).
- Franks, N. P. & Lieb, W. R. *J. molec. Biol.* **133**, 469–500 (1979).
- Lea, E. J. A. *Int. J. biol. Macromolecules* **1**, 185–187 (1979).
- Simon, S. A., Stone, W. L. & Busto-Latorre, P. *Biochim. biophys. Acta* **468**, 378–388 (1977).
- Gruen, D. W. R. & Haydon, D. A. *Biophys. J.* **30**, 129–136 (1980).
- Schoenborn, B. P. *Brookhaven Symp. Biol.* **27**, 1–10–1–17 (1975).
- Blasie, J. K., Schoenborn, B. P. & Zaccai, G. *Brookhaven Symp. Biol.* **27**, III-58–III-67 (1975).
- Zaccai, G., Blasie, J. K. & Schoenborn, B. P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 376–380 (1975).
- Worcester, D. L. & Franks, N. P. *J. molec. Biol.* **100**, 359–378 (1976).
- Büldt, G., Gally, H. U., Seelig, A., Seelig, J. & Zaccai, G. *Nature* **271**, 182–184 (1978).
- Büldt, G., Gally, H. U., Seelig, J. & Zaccai, G. *J. molec. Biol.* **134**, 673–691 (1979).
- Zaccai, G., Büldt, G., Seelig, A. & Seelig, J. *J. molec. Biol.* **134**, 693–706 (1979).
- McGlashan, M. L. & Williamson, A. G. *Trans. Faraday Soc.* **57**, 588–600 (1961).
- Fuhrer, H. *Biochim. Z.* **115**, 235–261 (1921).
- White, S. H. *Biophys. J.* **25**, 9a (1979).
- Worthington, C. R., King, G. I. & McIntosh, T. J. *Biophys. J.* **13**, 480–494 (1973).

MATTERS ARISING

Galileo's observation of Neptune

KOWAL AND DRAKE¹ have found a record in Galileo's notebooks of Neptune's position—the most precise placing being that of 28 January 1613 (Besselian date 1613.08). They note, however, a disagreement of 1 arc min between the position observed (O) by Galileo and that predicted by modern calculation (C)—using the orbit of the Jet Propulsion Laboratory, Development Ephemeris no. 102, for the C. They suggest that the effect of a yet-undiscovered exterior planet may explain this discrepancy. The 1613 residuals are listed in Table 1. To evaluate conservatively

was made during his search for Jupiter satellites, so Jupiter was in his field of view. The account and drawing for his 28 January 1613 entry both note that Jupiter, star 'a' (SAO119234), and star 'b' (the planet Neptune) were in a straight line. Kowal and Drake note that the suspect modern ephemeris correctly predicts the same lineup.

This line cuts the ecliptic at a descending angle of 25.9°. (See Table 1 or Fig. 5 of ref. 1.) Thus, if the modern orbit is approximately correct and Galileo accidentally distorted his drawing's scale, then we would expect a false latitude residual $F\Delta\beta$ corresponding to the false longitude residual $F\Delta\lambda$ according to the relation:

$$F\Delta\beta/F\Delta\lambda = \tan(-25.9^\circ) \quad (1)$$

Table 1 Residuals (O–C) for 28 January 1613

Geocentric	RA	$\Delta\alpha$	–42 arc s	Dec	$\Delta\delta$	+47 arc s
Geocentric	long	$\Delta\lambda$	–57 arc s	lat	$\Delta\beta$	+26 arc s
Heliocentric	long	Δl	–56 arc s	lat	Δb	+26 arc s

the exterior-planet hypothesis, I have computed the strength of the 1613.08 perturbation on Neptune, caused by hypothetical tenth planets in those circular orbits^{2,3} permitted by modern residuals (see Table 2). The elements and masses used were those indicated by least-squares analyses^{2,3} of Neptune longitude residuals, post-discovery as well as M. Lalande's 8 and 10 May 1795 observations (discord, –7 arc s both nights)⁴.

Table 2 shows that: the magnitude of the disturbances produced by the hypothetical exterior planets is at best in the range 10–20 arc s which, even allowing for some mass uncertainty, disagrees with the 1613 residual being nearly 1 arc min; (2) the direction is negative in all the cases listed which agrees with Table 1.

The unlikelihood of a residual as large as 1 arc min is also supported by another line of reasoning. Galileo's observation

Setting $F\Delta\lambda$ equal to the geocentric longitude residual $\Delta\lambda$ of Table 1, we substitute this into equation (1) and find that the accidental scale-distortion hypothesis predicts a false longitude residual

$$F\Delta\beta = +28 \text{ arc s} \quad (2)$$

This is within 2 arc s of the observed geocentric latitude residual $\Delta\beta$ of +26 arc s (Table 1).

Kowal and Drake note that our present uncertainty about the motion of Neptune is due to the relative shortness of the planet's continuously observed arc (less than a single full revolution since the 1846 discovery). In other words, a sizeable error in the baseline of the planet's motion may be masked by an alien perturbation. (The baseline error is $A + Bt$ in equation (1) of ref. 3. A is the error in longitude at epoch, and B is the error in the mean motion.)

However, there is no such thing as a baseline error in latitude motion (no linear expression appears in equation (2) of ref. 3, compare with equation (1) of ref. 3). Thus, if the 1613.08 latitude residuals produced by the hypothetical disturbers of refs 2, 3 are computed, it turns out that they are merely of order of magnitude 1 arc s—a far cry from the observed residual of 26 arc s.

Thus Galileo's illustration of Star SAO119234 and Neptune probably has an erroneous scale. However, if Galileo was at least correct in picturing Neptune closer to the star than modern theory predicts, then a negative 10–20 arc s longitude residual would be consistent with perturbation by an outer planet itself reasonably consistent with later longitude residuals. Hence, the slight nonlinearity of Neptune, star, and Jupiter might not be recognized.

DENNIS RAWLINS

2636 Grand Avenue,
San Diego, California, USA

1. Kowal, C. & Drake, S. *Nature* **287**, 311–313 (1980).
2. Rawlins, D. & Hammerton, M. *Nature* **240**, 457 (1972).
3. Rawlins, D. & Hammerton, M. *Mon. Not. R. astr. Soc.* **162**, 261 (1973).
4. Rawlins, D. *Astr. J.* **75**, 856 (1970).

KOWAL AND DRAKE¹ established that Galileo observed the planet Neptune and actually noted its motion. However, I must comment on their statement that "we are left, then, with the strong possibility that the ephemeris of Neptune is in error by a significant amount . . .".

It would be virtually impossible for the ephemeris of Neptune to be in error by over 26 arc s in latitude (the component of the residual perpendicular to Neptune's orbital plane) without "the existence of an unknown perturbation". As good fortune would have it, modern observations (since 1910) cover just that part of the sky in which Galileo made his observation (the exact crossing of $\alpha = 12^h$ occurred in 1943). As such, the orbital plane of Neptune, especially in that part of the sky, is established to within a small fraction of 1 arc s.

We recently attempted to incorporate Galileo's observation into the JPL Planetary Ephemeris. Even with the unrealistically high observation weight of 1 arc s, the attempt proved to be entirely unsuccessful. It was possible to adjust the longitude somewhat, but the 26 arc s latitude residual remained.

On the other hand, a perturbation of 26 arc s in latitude was easily produced using numerical experiments with an additional planet of a few Earth masses, even when the closest approach to Neptune was as distant as 10 AU. However, such perturbations usually tend

Table 2 Perturbations from hypothetical planets

Disturber's distance (AU)	Disturber's mean long		Disturber's mass ($R_\oplus/206265$)	Longitude perturbations (arc s)		
	1973.0	1981.0		1613.08	1795.35	1980.0
75	300°	304°	3.95	–8.7	–3.4	+1.1
	310°	314°	4.80	–11.4	–5.6	+1.0
	320°	324°	4.11	–9.5	–5.6	+0.6
60	310°	316°	1.55	–8.4	–1.6	+0.8
	320°	326°	2.90	–18.5	–4.9	+1.1
	330°	336°	2.61	–18.5	–6.2	+0.7
50	340°	346°	1.98	–13.8	–5.7	+0.3
	340°	348°	2.32	–4.1	–5.7	+1.1
	350°	358°	1.74	–10.9	–6.3	+0.5
	0°	8°	1.24	–12.3	–5.7	+0.1
	10°	18°	0.95	–11.6	–5.1	–0.1

to be a full order of magnitude larger in longitude than in latitude, whereas the ratio required here is only about 5:2.

To explain Galileo's residual as an ephemeris error, an unknown and rather pathological perturbation must be invoked. As Kowal and Drake state, "...it is surprising that the 'incorrect' ephemeris position still lies on the same straight line from Jupiter and the star".

One must seriously question Galileo's drawing as (1) Galileo was not interested in "extraneous objects", and (2) Galileo used a dashed line to connect stars 'a' and 'b' as he did in other cases where he wished to indicate direction only, not distance.

It is true that other pre-1900 observations of Uranus and Neptune do not seem to be consistent with those of the modern era. However, these discrepancies could be the result of observational errors, as most probably is the case here with Galileo.

E. MYLES STANDISH JR
Planetary Ephemeris Development,
Jet Propulsion Laboratory,
Pasadena, California 91125, USA

1. Kowal, C. T. & Drake, S. *Nature* **287**, 311-313 (1980).

KOWAL REPLIES—Although I am pleased to see these responses to our letter, I feel that an adequate assessment of Galileo's accuracy requires a much more rigorous analysis. In particular, Rawlins' assumption of a circular orbit for an 'outside' perturber is likely to be a gross oversimplification, as it would be in the case of Pluto.

Standish's reasons for questioning Galileo's drawing are not valid, for the following reasons:

Galileo clearly was interested in this "extraneous object" because he specifically noted its relative motion. Furthermore, he drew a linear scale in that diagram for the very first time. It would be quite surprising if he did not use that scale when drawing Neptune's position.

Galileo often used dashed lines, even when indicating exact distances; for example, in the drawings of SAO119234 on 25-27 January. It is certainly true, however, that Galileo's drawings are far less accurate than his measurements.

Standish's statement that the requisite perturbation in latitude was "easily produced" in his numerical experiments is suggestive. Again, Pluto shows that "pathological" cases do exist.

In any case, I hope that the question of Galileo's accuracy will not obscure the central points of our paper. These are:

We have established that Galileo observed Neptune, and noted its motion, 234 yr before Neptune was discovered.

We have demonstrated a new technique for finding other pre-discovery observations of Neptune. In particular, we should search for observations of the 1702 occultation.

Obviously, I would not insist that Galileo's drawing is accurate. If more pre-discovery observations of Neptune are found, we will be able to make a more useful assessment of the Neptune residuals. Until then, I must maintain that there is a "strong possibility" that the theory of Neptune's orbit needs to be revised; not only because of the Galileo observation, but also because of the previously known residuals.

CHARLES T. KOWAL
Department of Astrophysics,
California Institute of Technology,
Pasadena, California 91103, USA

Does spider feeding behaviour optimize dietary essential amino acid composition?

GREENSTONE¹ has concluded that the spider *Pardosa ramulosa* takes a non-random mixture of its three main prey and that this mixture optimizes the proportions of the essential amino acids in the diet. For statistical reasons, these conclusions are invalid.

The spiders fall into eight classes, which Greenstone regarded as the cells of a $2 \times 2 \times 2$ contingency table, according to whether they were serologically positive or negative for each of the three main prey. The overall proportion of triple negatives is about 25%, from which 3% must be (and was) removed to allow for moulting animals, which do not feed. If the average number of serologically potent prey in the gut is one, the remaining proportion of triple negatives is consistent with Greenstone's view that about 20% of the diet is made up of species other than the three main ones. However, some guts contain at least three prey individuals, as they are positive for all three species. Thus, there must be considerable spread around the mean, with many spiders containing no prey and some containing many.

As a result, the contingency tables may show large values of ' χ^2 ' even if the recent occurrence of one species in the diet has no effect on the relative probabilities of the spider eating another member of the same species or a member of another species.

This may be demonstrated by a simple example. Suppose we have a population comprising N_0 animals with no prey individuals in the gut, N_1 with one and N_2 with two. Further suppose that the prey are of two species, which occur in the overall proportions of p and q in the diet. Finally, suppose that these proportions are the same among the animals with one prey individual in the gut as among those with two prey individuals in the gut. Among the latter, if the presence of one species in the gut is unrelated to the presence of the other, then the propor-

tions of individuals with two individuals of the same species in the gut are p^2 and q^2 ; a proportion $2pq$ will have one prey of each species in the gut. Thus, in the population as a whole, there would be the following numbers of individuals serologically positive or negative for the two species:

$$A-, B-: a = N_0$$

$$A+, B-: b = pN_1 + p^2N_2$$

$$A-, B+: c = qN_1 + q^2N_2$$

$$A+, B+: d = 2pqN_2$$

If we treat these four classes as the cells of a 2×2 contingency table, our 'expectation' is that $ad - bc = 0$. Except for particular sets of parameter values this expectation is clearly false. In other words, the calculation of χ^2 does not provide a test of the hypothesis that the probability of consumption of each species is independent of whether or not the spider has recently consumed a member of that species.

An alternative explanation of the high proportion of triple negatives is that the proportion of the diet made up by species other than the main three is higher than 20%. If so, it scarcely seems right to ignore these other species when working out the dietary composition.

J. J. D. GREENWOOD
Department of Biological Sciences,
University of Dundee,
Dundee DD1 4HN, UK

1. Greenstone, M. H. *Nature* **282**, 501-503 (1979).

GREENSTONE REPLIES—In all phases of human endeavour it is important to mind one's p's and q's; in statistics one must also ask whether they are relevant to the question at hand. Greenwood's exposition of the prey item makeup of the serologically positive stomach classes is correct and self-evident. However, only the values of a , b , c and d , and not their underlying constituents, are relevant to the test of association.

Let us consider a hypothetical human dietary example. Suppose a cafeteria serves only beans and rice and gives each diner the choice of one, two or no scoops of each. After a large sample of diners has passed through the line we tabulate the contents of their plates. If we wish to know whether the average diner is receiving a balanced diet we must add up all of the scoops of beans and rice so as to compute Greenwood's p and q , and then apply them to what we know about human nutritional requirements and the nutrient compositions of beans and rice; this is analogous to my use of the data in Tables 2 and 3 of the original paper¹ to compute the proportions of the essential amino acids in the spiders' diets.

If, however, we wish to know whether beans and rice are significantly associated in the diet, we need only concern

ourselves with the numbers of empty plates, of plates containing no rice but some (that is, at least one scoop of) beans, of plates containing no beans but some rice, and of plates containing some of each. These numbers are a , b , c and d of our contingency table, and we use the marginal totals to compute the expected numbers of each of the four kinds of plates. The numbers of scoops of beans and of rice on plates having at least some of either are no more necessary to this test of association than are the severity of cancer or the number of packs smoked per day to a test of association between smoking and lung cancer.

MATTHEW H. GREENSTONE

Department of Entomology
and Nematology,
University of Florida,
Gainesville, Florida 32611, USA

1. Greenstone, M. H. *Nature* **282**, 501–503 (1979).

Plasmoid confinement by the charged particle micro-fields

DIJKHUIS¹ has recently proposed a 'boson model' which explains ball lightning in terms of electric charge. It is based on the hypothesis that turbulence in the ball lightning plasma enhances charge transfer to the point that charged vortices transport electricity the way electrons do in a superconductor. The result is described as a sponge-like structure of tenuous vortex cores nested in dense ambient plasma. By reasonings based on Landau damping and local deviations from charge neutrality it is argued that the size of the plasma vortices behaving as charged bosons apparently lies around 1 Debye length.

Plasmoid models of ball lightning are normally taken to be ruled out by the plasma virial theorem yielding the result, at least in its usual continuum formulation, that a plasmoid—defined as a self-contained bundle of electromagnetic and material energy—cannot exist. The theorem states that the plasmoid energy contents—electromagnetic as well as gas kinetic—are all positive and together drive expansion. However, the boson model for plasma turbulence is claimed to lift the virial constraint on maximum energy content with negative potential energy from interaction between plasma vortices.

The negative interaction energy necessary for the boson model conflicts with the positive-definite nature of all the energies in the fluid virial theorem. Note that the proposed boson model specifically involves interactions and structures on the plasma microfield scale that is, the Debye length. As discussed by Gerjoy and Stabler² the averaged, 'smoothed-out', plasma descriptor inherent in the fluid virial theorem may, in some respects, be misleading. Resort has then to be taken in

a perfectly general and purely particle system formulation of the virial theorem. The electrostatic energy is then given by the double sum

$$W_{el} = \frac{1}{4\pi\epsilon_0} \sum_{i < k} e_i e_k r_{ik}^{-1}; r_{ik} = |\mathbf{r}_i - \mathbf{r}_k| \quad (1)$$

Note that a slight excess of negative charges around each positive charge will make W_{el} go negative². However, this is also that condition which will distribute the source charges e_k so as to minimize the electrostatic potential ϕ^i at the field point i as well as to make the related field strength $\mathbf{E}^i = -\nabla\phi^i$ small

$$\phi^i = \frac{1}{4\pi\epsilon_0} \sum_{k \neq i} e_k r_{ik}^{-1} \quad (2)$$

The influence of a negative contribution W_{el} , as indicated above, to a virial theorem reasoning has been discussed and estimated in ref. 2. For a plasma with thermal energy like that assumed for the boson model it has been proved to be insignificant, thus validating the results of simpler fluid virial theorem reasoning.

In a similar way, the magnetic energy W_m and vector potential $\mathbf{A}^i$ can be expressed in terms of the microscopic currents $\mathbf{j}_i$ and $\mathbf{j}_k$ carried by the charges e_i and e_k

$$W_m = \frac{\mu_0}{4\pi} \sum_{i < k} \mathbf{j}_i \cdot \mathbf{j}_k r_{ik}^{-1}, \mathbf{A}^i = \frac{\mu_0}{4\pi} \sum_{k \neq i} \mathbf{j}_k r_{ik}^{-1} \quad (3)$$

Here it is noted that a hypothetical current distribution where adjacent current elements are directed oppositely would make W_m go negative but, also, make $\mathbf{A}^i$ and $\mathbf{B}^i = \text{curl } \mathbf{A}^i$ attain vanishingly small values. Contrary to the claim in ref. 1, there seems to be no way by which the magnetic interaction energy can attain a large enough negative value to affect significantly the expansion-driving thermal effects. Essentially the same result, in conflict with the proposed boson model for ball lightning, was also obtained by Liboff and Lie³ who used instead the quasi-relativistic Coulomb gauge expression for $\mathbf{A}^i$ in their derivation of a general virial theorem in a particle formulation.

I conclude that the various formulations of the virial theorem as discussed above are all based on usual charged-particle electrodynamics where fields and forces follow from positions and velocities. The power then of the theorem to deny plasmoid self-confinement lends support to that theory⁴ where the confinement is provided by the radiation fields associated with particle acceleration.

E. A. WITALIS

Research Institute of National Defense,
Stockholm, Sweden

DIJKHUIS REPLIES—It is a delusion to claim perfect generality for the purely particle system formulation of the virial theorem. Particles in a plasma obey classical equations of motion as long as electromagnetic fields do not change appreciably over their dimension. This classical limit is certainly appropriate for any stable plasma in a smooth electromagnetic field. However, in a turbulent plasma electromagnetic fields may have strong local fluctuations. In our approach to ball lightning, turbulent fluctuations of local electromagnetic fields become so violent that the motion of particles must be treated quantum mechanically.

A particle with mass m and charge q , moving with velocity $\mathbf{v}$ in an electromagnetic field with vector potential $\mathbf{A}$ and electrostatic potential ϕ , has canonical momentum $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$ and hamiltonian $H = (\mathbf{p} - q\mathbf{A})^2/2m + q\phi$. In our quantum-mechanical context, $p = -i\hbar\nabla$ is an operator which does not commute with arbitrary functions of position vector $\mathbf{r}$. Thus $\frac{1}{2}q(\mathbf{v} \times \mathbf{B} - \mathbf{B} \times \mathbf{v})$ (ref. 1) denotes the Lorentz force on a wave packet moving in a magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$. In the absence of magnetic fields a wave packet has $\mathbf{p} = m\mathbf{v}$, $H = \frac{1}{2}m\mathbf{v}^2 + q\phi$, and the virial theorem reads $d/dt \langle \mathbf{r} \cdot m\mathbf{v} \rangle = m\langle \mathbf{v}^2 \rangle - q\langle \mathbf{r} \cdot \nabla \phi \rangle$ (ref. 1), where the brackets denote expectation values with wave functions from the energy representation. Inclusion of magnetic interaction between wave packets leads after lengthy commutator bracket calculations to a quantum-mechanical virial theorem of the form

$$\frac{d}{dt} \langle \mathbf{r} \cdot m\mathbf{v} \rangle = m\langle \mathbf{v}^2 \rangle + \frac{1}{2}q\langle \mathbf{r} \times \mathbf{v} \cdot \mathbf{B} + \mathbf{B} \cdot \mathbf{r} \times \mathbf{v} \rangle - q\langle \mathbf{r} \cdot \nabla \phi \rangle \quad (1)$$

where for a stable plasmoid at least one term on the right has to be negative. With $\rho^{1/2}e^{i\theta}$ as wave function for a boson state with charge density ρ and phase θ (ref. 2), expectation values in equation (1) include the quantum-mechanical term $-\hbar^2/2m(\nabla^2\rho^{1/2})/\rho^{1/2}$, from which a spherically symmetric solution with negative energy was derived in our letter³. In the classical limit $\hbar \rightarrow 0$, expectation values can be dropped and all quantities commute, so that the usual terms for kinetic, magnetic and electric energy are recovered.

I maintain that (1) the quantum-mechanical version of the virial theorem admits electromagnetic self-confinement of a plasma, and (2) strong plasma turbulence leads to a self-confined boson state in ball lightning.

G. C. DIJKHUIS

Zeldenrust College, Terneuzen,
The Netherlands

1. Schiff, L. I. *Quantum Mechanics* 3rd edn, 177–180 (McGraw-Hill, New York, 1968).
2. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* Vol. III, Section 21 (Addison-Wesley, New York, 1966).
3. Dijkhuis, G. C. *Nature* **284**, 150 (1980).

1. Dijkhuis, G. C. *Nature* **284**, 150 (1980).
2. Gerjoy, S. & Stabler, R. C. *Phys. Fluids* **7**, 920 (1964).
3. Liboff, R. L. & Lie, T.-J. *Phys. Fluids* **11**, 1943 (1968).
4. Bergström, A. *Phys. Rev. D* **8**, 4394 (1973).

BOOK REVIEWS

Science and faith in mind

Stuart Sutherland

Brains, Machines and Persons. By Donald MacKay. Pp.114. (Collins/W. Eerdmans: 1980.) Pbk £3.50, \$4.95.

BOTH as a brain scientist and as a Protestant, Professor Donald MacKay is an honest man. *Brains, Machines and Persons* gives a brief but stimulating and lucid account of the human brain and of models of its workings cast in the form of computer programs. This account is a preamble to an attempt to reconcile the scientific and the religious view of human beings. He makes a better shot at it than others, such as Sir John Eccles, who have confronted the same dilemma. Professor MacKay rightly eschews both the argument for free will based on the Heisenberg uncertainty principle and any form of interactionism between mind and body, though he does not set out the reasons why these proposals are unsatisfactory. (The first can only account for random behaviour whereas the outcome of deliberate choice is far from random; the second implies that the brain is not a physical system in the ordinary sense of the word since at some point it is being influenced by non-physical events.)

MacKay's own solution to the mind-body problem is that human beings can be viewed in two ways, from the inside as a series of mental processes (the I-story) and from the outside as a series of physical events. He does not deny that the workings of the brain and nervous system may be deterministic and, at least in principle, predictable. He resolves the problem of freedom of the will by arguing that if someone predicts from a knowledge of his own brain state that he will do something, his brain is itself changed by this knowledge: this change may lead to the prediction being falsified since the brain cannot take into account the effects of the change in its own state in order to predict its future behaviour without falling into an infinite regress. MacKay then raises the question "suppose a super-scientist... produced a specification of [your brain's] immediate future... Would you be correct to believe it?". He argues that you would not, since when the super-scientist gave you his prediction, your brain state would change and hence his prediction might fail. This is an argument with which I have dealt elsewhere ("Is the Brain a Physical System" in *Explanations in the Behavioral Sciences* edited by R. Borger and F. Cioffi, Cambridge University Press; 1970). The problem is that the super-scientist might have taken into his own calculations the effects of telling you what

you will do, and have delivered the prediction only after making sure that you will still do what he predicts even after being informed that you will do it. MacKay is right in thinking that if the brain scientist delivers *all* the thought processes that underlie his prediction, then your brain state may be changed and the prediction may fail: once again an infinite regress arises. But MacKay fails to note that you may have other good reasons for believing the super-scientist, for example, you may know that he always gets his predictions right.

Despite this flaw in MacKay's argument, there is something to be said for the theses that mental events and physical states of the brain are one and the same thing viewed in different ways and that we can only choose freely because we cannot ourselves predict what we will do. He argues that it is the latter aspect of people that makes each a morally responsible agent. Even if this were true, it leaves open the question of how we view others. If they can legitimately be viewed as complex physical systems, does it make sense to praise and blame them? To the extent that we know the cause of an action, we tend not to attribute praise or blame. Someone who becomes extremely bad-tempered because of a tumour in the hypothalamus does not deserve censure. Here MacKay draws a distinction between brain hardware and brain

software. If the hardware has gone wrong, no moral responsibility is attached. If it is the program being executed in the brain that is at fault, then it is proper to attach blame. Ingenious though this argument is, it again seems fallacious: suppose we knew enough about someone to predict from his genetic make-up and his environmental history exactly what programs were running in his brain, would that not make us want to excuse him for evil acts?

MacKay makes further play with the distinction between the brain's hardware and software. He even suggests that one solution to the problem of resurrection would be to incorporate the program that runs in each individual's brain in some medium other than the human body. He faces the fact that in the distant future we may be able to build computers as intelligent as ourselves and having many, if not all, of the attributes that we value in ourselves (other than being born of man). He believes that the existence of such machines would not lower the status of human beings, though it would raise that of computers. I once asked him whether such a computer would be a candidate for entry into the kingdom of heaven: with his usual honesty he replied, after careful reflection, that he could see no reason why not. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.

Activity in the blood-brain barrier

Michael Bradbury

The Cerebral Microvasculature: Investigation of the Blood-Brain Barrier. Edited by Howard M. Eisenberg and Robert L. Suddith. Pp.342. (Plenum: 1980.) £24.89, \$39.50.

AFTER the idea of a blood-brain barrier arose at the turn of the century, there was prolonged controversy as to whether it was sited at the level of the capillary endothelium or at that of the sheath of glial end-feet. In 1946, Krogh showed great foresight in not only placing the barrier at the capillary level, but in considering it to have the typical properties of a cell membrane and perhaps to be capable of secretion, i.e. active transport. Physiological and ultrastructural studies in the whole animal have fully confirmed this interpretation, and it is now accepted that

the cerebral endothelium functions like a tight epithelium and is capable of active transport.

The Cerebral Microvasculature contains the edited manuscripts from a symposium held at Galveston, Texas, in 1979. The first section deals with transport properties and permeability. It largely contains reports of the properties of isolated whole capillaries of cultured endothelial cells and of parts of capillaries, plasma and basement membranes for example. Among a number of interesting papers, the uptake of various solutes into isolated capillaries as a model for transport across the barrier *in vivo* is reviewed by Betz and Goldstein, and the biochemical properties of isolated capillaries are considered by Mršulja and Djuričić. Eisenberg and his colleagues discuss the demonstration of an ouabain-

sensitive potassium pump at the abluminal membrane of the endothelial cells and relate this to a high concentration of the Na-K-ATPase in these cells compared with that in both non-cerebral endothelial cells and in cerebral cortex as a whole. Brendel and Meezan look critically at the technique for isolating cerebral capillaries and compare the composition of the basement membrane from brain vessels with that from retinal vessels and renal glomeruli.

One is more convinced of the potential of isolated capillaries and cultured endothelial cells as a means of studying control of transport at the blood-brain barrier than as a model for transport across it. Thus, DeBault and Cancilla report intriguing experiments in which the co-culture of endothelial cells with glioma cells leads to a maintenance of a high content of γ -glutamyl transpeptidase in the former. This enzyme is highly specific to cerebral capillaries and the results strengthen the concept that the glial end-feet induce the very particular properties of these vessels. In the same context, Anders and colleagues describe curious "assemblies" of intramembranous particles which occupy the cytoplasmic face of astrocytic plasma membranes. They are particularly numerous in the perivascular membranes of the glial end-feet and could be part of a mechanism for influencing the capillaries.

Isolated capillaries contain adenylyl cyclase (as discussed by Palmer and colleagues) and this enzyme is activated by noradrenalin (mainly β_2 effect), dopamine and in the rabbit by histamine. This possible link in a control of permeability is of particular interest, since cerebral capillaries are contacted by noradrenergic nerve terminals probably originating from cell bodies in the locus coeruleus. The evidence that these control permeability to labelled water independently of any effect on cerebral blood-flow is considered by Hartman and colleagues.

The second section of the book is concerned with perturbation of the barrier and lacks, for me at least, the excitement of the first section. It contains good articles by Nemoto on transport during anaesthesia, by Rapoport on measurement by his new technique of permeability in relation to lipid solubility and of permeability after osmotic shock, and a definitive short review of the blood-brain barrier in acute and chronic hypertension by Johansson.

This book is much better than the usual collection of manuscripts from a symposium. It should be of interest to all those concerned with the blood-brain barrier or with transport across the capillary wall in general. Most particularly, it is valuable as a fairly comprehensive survey of the present state of the art with regard to isolated cerebral capillaries and endothelial cells. □

Michael Bradbury is Professor of Physiology at King's College, University of London, and author of *The Concept of the Blood-Brain Barrier* (Wiley, 1979).

American Ice Age mammals

R.J.G. Savage

Pleistocene Mammals of North America. By Björn Kurtén and Elaine Anderson. Pp.442. (Columbia University Press: 1980.) \$42.50, £21.20.

BJÖRN Kurtén's *Pleistocene Mammals of Europe* (Weidenfeld & Nicolson, 1968) has become a classic; it broke new ground and remains a frequently quoted text. Now the author, in collaboration with Elaine Anderson, presents the *Pleistocene Mammals of North America*. The format of the new book is similar, but more detailed treatment is given to major sites and their faunas are compared with those of adjacent continents, a useful feature lacking in the earlier work. Although entitled "Pleistocene" mammals, the authors wisely extend their coverage back to include the whole of the Blancan, to a base of around 3.5 million years. Their geographical limits are Canada, the USA and northern Mexico.

In Part I some 250 faunal sites are listed, with notes on geology, geography, ecology and faunal composition. The listing is alphabetical for each state, and the faunas are grouped as Blancan, Irvingtonian and Rancholabrean. There follows a discussion of intercontinental correlation and migrations between North and South America and between North America and Eurasia.

In the second part of the book critical and annotated lists are given for all Blancan and Pleistocene mammal species on the continent. The taxonomic treatment is detailed, giving information on synonymy, geographical and stratigraphical distribution, anatomical characteristics, relationships, evolution, migration and extinction of each species. Some 562 species are treated; extant species are included only where they have a fossil record. The diagnostic features detailed for each species of necessity use highly specialized terms, which however are not explained; e.g. on p.285 a horse species is characterized by having "ectoflexid of lower molars penetrating deeply into the metaconid-metastylid isthmus". In this second section the material ancillary to the text leaves something to be desired. The illustrations depict only a few sample species and the dental characters are not annotated. The quality of the drawings is highly variable; most are taken from the literature without alteration, thus the orientations and scales are often inconsistent and the techniques of drawing differ widely. Further, they are poorly arranged so that there is much wasted space. The maps showing species distribution are clear, though not numerous, and the addition of some reconstructions adds interest, especially for the non-specialist. The tables are important; in Appendix 2 the stratigraphical range of every species is tabulated.

The book is a formidable and scholarly compendium of information: thorough, accurate and up to date. It will be an invaluable reference source for those working not only on Pleistocene faunas, but on problems of evolution, extinction and biogeography. The authors are indeed to be congratulated on so ably treating a colossal mass of data in a digestible and attractive form. □

R.J.G. Savage is Reader in Vertebrate Palaeontology at the University of Bristol.

Stone Age summary

Derek Roe

Early Man in Britain and Ireland: An Introduction to Palaeolithic and Mesolithic Cultures. By Alex Morrison. Pp.209. (Croom Helm/St Martin's Press: 1980.) Hbk £12.95, \$35; pbk £7.95.

THE Mesolithic period in Britain lasts for around 5,000 years, the Palaeolithic for perhaps 500,000. Books on British archaeology seem to be produced in an inverse relationship to the durations of the successive periods. Lavishly illustrated information on Saxons, Vikings, Romans or Iron Age groups beckons from every shelf, but the choice is limited for those trying to study Early Man in Britain. Clearly, Alex Morrison's book meets a real need and let us hope the reading public will think so as well as teachers and students of archaeology, because early Prehistory in Britain, however unspectacular by comparison with, say, its equivalent in East Africa, has considerable fascination and importance in an Old World context.

That, however, does not make it an easy subject for a book. Sites exist in thousands, but the material from them is widely scattered; dating, for the Palaeolithic, is extremely difficult, and relatively few fully documented modern excavations have taken place. Undisturbed sites are rare: usually all fragile objects and much of the environmental evidence has gone. The British Mesolithic is a little better off in these respects, but even here the situation is far from ideal. For broad interpretation, one needs to know much about the Palaeolithic and Mesolithic of continental Europe: the geographical situation of Britain positively demands this, as Mr Morrison rightly stresses.

In these circumstances, one can either set out to write a massive book on the Palaeolithic and another on the Mesolithic, each based on exhaustive first-hand study of the extant material, or one can produce an outline summary, based almost entirely on the literature. Alex Morrison has opted

for the latter, and makes no extravagant claims for the result. Specialist colleagues will doubtless complain of brevity, unevenness, generalizations and omissions, but such specialists are relatively few and instead of complaining they should write the definitive texts whose absence they deplore.

In his two opening chapters, the author covers essential introductory subjects, including Pleistocene geology and environments, long-range dating methods, hominid evolution and so forth. Palaeolithic archaeology is very different from other kinds and preliminary discussion of these topics is certainly needed. The rest of the book mainly follows the archaeological sequence from Early Acheulian and Clactonian to the end of the Mesolithic, with an intermission to discuss the environmental changes that mark the start of the "postglacial" period, as we call it somewhat naively, since it's really only another interglacial. The approach depends heavily on individual British sites throughout, and the Mesolithic is studied regionally. Plenty of references are given, which serious students will need to follow up to supplement the brief descriptions provided. Some major relevant sites in Europe, and occasionally further afield, are mentioned in passing, though there is no serious attempt to set the British sequence firmly in its context stage by stage.

Plenty of major questions arise as soon as one starts to study the British material. For example, how does the British local Pleistocene sequence correlate with various fuller and better documented ones in Continental Europe? How do the named stages in Britain match the numbered "isotope stages" of the deep-sea core record? There are clearly gaps in the British succession, but just where do they come? Are we at present failing to observe a whole glacial-interglacial cycle, or even more, somewhere in the British Middle Pleistocene, thereby misleadingly con-

tracting our geological and archaeological sequences? How early is our "Early Acheulian" by world standards? What role does the Clactonian play? Why is the British Middle Palaeolithic so sparse, when the French Mousterian is so rich and varied? Has our Upper Palaeolithic anything at all to do with the classic sequence in southwest France, or can the whole of it be explained better by reference to central and northwest Europe? How much of our Mesolithic is attributable to direct connections with the Continent across a North Sea land-bridge, and how much represents indigenous development within Britain, showing variability related to the exploitation of differing environments and local resources? These are a few of the crucial and fascinating problems still to be solved: the reader will find awareness of most of them in the book, but might have liked more guidance for his thinking about possible solutions.

One never knows, these days, whether illustrations reflect an author's approach or a publisher's budget. Here, they are of rather mixed quality and usefulness. The line drawings are clear, but many of the photographs contrive to conceal as much information as they provide. All the photographs are of artefacts, chosen entirely from local sources in Glasgow, where the author is a lecturer in archaeology at the university: in a few cases are they the best examples available in Britain to document the point in question and some are even unprovenanced finds. The drawings include several maps, but hardly any plans or drawn sections of sites. The text of the book, which contains much that is up to date, deserved more generous and varied illustration, but perhaps one kind of expansion would have led to another and it would have ceased to be an outline account at all. □

Derek Roe is University Lecturer in Prehistoric Archaeology and Director of the Donald Baden-Powell Quaternary Research Centre at Oxford.

Polymer trio

Herbert Morawetz

Polymers, in 3 parts. *Methods of Experimental Physics*, Vol.16. Edited by R.A. Fava. Part A pp.590; Part B pp.421; Part C pp.528. (Academic: 1980.) \$55, £30.80; \$45, £25.20; \$59, £33.20.

THE publication of an encyclopaedic treatise with the scope of these three volumes is a major event for polymer physicists. As with all monographs in which chapters are written by experts, the success of the enterprise depends to a large degree on the manner in which the editor has defined the intended readership, selected the areas of greatest interest and chosen the authors who could best survey these areas. The first criterion does not appear to have been always considered; thus, while most of the chapters demand of the reader a high degree of scientific sophistication, several are written at a much lower level. Further, some contributors discuss almost exclusively experimental techniques, while others include valuable examples of experimental results. Most of the authors were well chosen and some may be regarded as leading authorities in their field. As for the choice of subject matter, it would seem to me that some contributions — "Inelastic Electron Tunneling Spectroscopy" and "Positron Annihilation", for example — are of rather marginal interest to the polymer physicist, while the absence of chapters on dynamic light scattering and on ESCA is to be regretted.

The first of the three volumes covers molecular structure and dynamics. For me, the highlight is the superb chapter on high-resolution NMR spectroscopy. It covers the numerous applications very well and includes broad-line NMR studies. My only regret is the absence of reference to the polymer association studies of Schneider and Spevacek and a mention of NMR studies of helix-coil transitions in polypeptides. Other outstanding chapters are those on infrared and Raman spectra, Rayleigh-Brillouin scattering in polymers and small-angle light scattering.

It is unfortunate that this valuable volume is prefaced by a rambling introductory chapter which falls far short of what is required. The chapter on molecular weights is extremely poor and contains many glaring errors. It is also a pity that the volume contains no mention of viscoelasticity, dielectric dispersion and ultrasonics (discussed in Part C for polymers in bulk) which have important applications to polymer solution studies.

Part B is subtitled "Crystal Structure and Morphology". Here, my major criticism is that several subjects of considerable interest are missing. There is no mention of the macroscopic crystals which can be obtained from diacetylene derivatives in a process in which the



By Alan Sorrell, from *Early Man in Britain and Ireland*.

Reconstruction of activities at Starr Carr, Vale of Pickering, Yorkshire, c. 7,500 BC.

polymer grows homogeneously in the monomer crystal. Epitactic polymerization is another phenomenon which might have been mentioned, and I feel that electron diffraction should have been treated more fully.

In Part C ("Physical Properties") the authors have apparently assumed that they were to write about polymers in bulk. This is an unfortunate limitation in an otherwise excellent volume, since a number of the techniques yield valuable data when applied to dilute polymer solutions. Plazek's excellent chapter on viscoelastic and steady-state rheological response would have benefited from the inclusion of experimental data; it is most surprising that the dependence of polymer viscosity on chain length and temperature is not mentioned. Also, Rutherford and Brown might have discussed the effect of the rate

of loading on stress-strain measurements and Hay's treatment of polymer orientation could have included a mention of anisotropic polymer solutions. The chapter on polymer alloys stresses excessively cohesive energy density, which is known to be inadequate in predicting polymer compatibility. As for the chapters on the electrical behaviour of polymers, I wish that polymer conductivity had received a fuller treatment and that the important phenomenon of piezoelectricity had been at least mentioned.

However, the various criticisms detailed above should in no way detract from the value of this survey. The books should be part of the library of every serious polymer scientist. □

Herbert Morawetz is Director of the Polymer Research Institute at the Polytechnic Institute of New York.

The philosophy of geochemistry

Arthur W. Rose

Environmental Geochemistry: A Holistic Approach. Ecological Studies, Vol. 35. By John A.C. Fortescue. Pp. 347. (Springer-Verlag: 1980.) DM 69, \$40.80.

ALTHOUGH this book is entitled *Environmental Geochemistry*, its main emphasis is on landscape geochemistry and related aspects of geochemistry in the surficial zone of the Earth. Landscape geochemistry deals with the chemical interactions between rocks, the atmosphere, natural waters and living organisms, and thus furnishes the geochemical background for many environmental problems. Most previous discussion of the subject has been by Russians, notably A.I. Perel'man. A major contribution of this book is to summarize the Russian approach and results, which are referred to throughout the text.

The opening chapters describe landscape geochemistry in broad outline and also the purpose of the book, which is to introduce a general approach and mode of thinking for conducting environmental geochemical studies and for comparing results from different areas; the history and concepts of landscape geochemistry are summarized (geochemical cycles, landscape prisms, landscape types, mobility of elements), together with the permutations and complications created by space, time, chemical complexity and varying scientific effort.

Each of the important concepts and principles of landscape geochemistry is then discussed in detail. These include element abundance in rocks, waters, plants, animals and air, and the speciation of these elements into abundances of different form (exchangeable, acid extractable and so on); element migration

or relative mobility of elements under varying conditions; geochemical flows by various types of surface and near-surface transport; geochemical gradients such as decrease of chemical content away from an orebody or pollution source, and climatically controlled gradients; geochemical barriers formed by changes in chemical environment along a flow path, resulting in precipitation and accumulation of selected elements; historical geochemistry, leading to recognition of changes in chemistry with time; and geochemical classification of landscapes. These chapters include many tables giving basic geochemical data and numerous examples and figures from environmental and exploration geochemical studies.

The book concludes with an interesting chapter on practical applications of landscape geochemistry which illustrates uses in geochemical exploration for minerals, environmental pollution, waste disposal, plant and animal nutrition, and geoepidemiology. An integrated view of all elements in a landscape prism and of changes with time and space is stressed throughout.

The author has a philosophical and pedagogical approach to his subject, and repeatedly emphasizes the organization of ideas and information rather than the information itself. Most experienced environmental and exploration geochemists are probably familiar with many or all of the concepts discussed, but the book draws from a diverse background, and all those investigating chemical processes on the Earth's surface will probably gain new ideas from the book. □

Arthur W. Rose is Professor of Geochemistry at The Pennsylvania State University and is co-author of Geochemistry in Mineral Exploration (Academic Press, 1979).

Taxonomy is far from boring

Otto Kandler

Microbiological Classification and Identification. The Society for Applied Bacteriology Symposium Series, No 8. Edited by M. Goodfellow and R. G. Board. Pp. 408. (Academic: 1980.) £20, \$46.

CLASSIFICATION and identification is often considered to be a boring, old-fashioned subject dealing with the simple morphological and physiological properties of organisms. This volume, however, demonstrates the opposite. The 15 papers, presented by leading experts at the 1978 SAB symposium at Birmingham, describe the highly sophisticated applications of the most advanced biochemical and molecular biological principles and techniques for exploring and determining new taxonomical characters. They do not, however, give detailed prescriptions on how to carry out the identification of a particular organism, nor do they contain phylogenetic trees or keys for the classification of particular taxa or groups of bacteria. Since none of the contributions cover viruses or fungi, the reader may feel somewhat misled by the term "microbiological" used in the title.

One paper deals with the principles of computer-aided numerical taxonomy in a way quite understandable to the non-mathematician. Others are concerned with the application of DNA-DNA hybridization and with the role of neutral mutations, insertion sequence elements and plasmids in the genotypic variability of bacteria. Aspects of fine-structure research and cell wall and lipid chemistry are also dealt with. In addition, papers are devoted to the applicability of enzyme and protein patterns of various kinds and to the chemical analysis of fermentation products. The suitability of serological methods is discussed with respect to their use in detecting food-poisoning toxins, the ability of phage-typing with respect to the classification of plant-pathogenic bacteria, and the usefulness of determining enzyme activities is exemplified by the identification of clinical specimens.

Apart from the omission of viruses and fungi, there are other gaps — I looked in vain for a concise contribution on the implications of sequence analysis of proteins and nucleic acids for the classification of taxa above the species level. Altogether, though, this volume will be most informative for advanced students and research workers in bacterial taxonomy; it will, however, not cover the everyday needs of the clinical or food bacteriologist. □

Otto Kandler is a Professor at the Botanisches Institut der Universität, Munich.

BOOKS RECEIVED

Mathematics

CHATFIELD, C. and COLLINS, A. J. *Introduction to Multivariate Analysis*. Pp.246. Hbk ISBN 0-412-16030-7; pbk ISBN 0-412-16040-4. (Chapman & Hall: 1980.) Hbk np; pbk £7.50.

HARRIS, C. J. and MILES, J. F. *Stability of Linear Systems: Some Aspects of Kinematic Similarity*. Mathematics in Science and Engineering, Vol. 153. Pp.236. ISBN 0-12-328250-0. (Academic: 1981.) £9.80, \$23.

MICHEL, A. N. and HERGET, C. J. *Mathematical Foundations in Engineering and Science: Algebra and Analysis*. Pp.484. ISBN 0-13-561035-4. (Prentice-Hall: 1981.) \$25.95.

MILLER, W. and WRATHALL, C. *Software for Roundoff Analysis of Matrix Algorithms*. Pp.151. ISBN 0-12-497250-0. (Academic: 1980.) \$18.50.

Astronomy

BÖHME, S. *et al.* (eds). *Astronomy and Astrophysics Abstracts*. A Publication of the Astronomisches Rechen-Institut Heidelberg, Vol. 27. Literature 1980, Part 1. Pp.939. ISBN 3-540-10479-8. (Springer Verlag: 1980.) DM 118, \$69.70.

GIACCONI, R. and SETTI, G. (eds). *X-Ray Astronomy*. Proceedings of the NATO Advanced Study Institute held at Erice, Sicily, July 1979. NATO Advanced Study Institutes Series C—Mathematical and Physical Sciences, Vol. 60. Pp.406. ISBN 90-277-1156-9. (Reidel: 1980.) Dfl. 90, \$47.50.

MITTON, J. A. *Language of its Own*. Key Definitions in Astronomy. Pp. 168. Hbk ISBN 0-584-10547-9; pbk ISBN 0-584-10565-7. (Frederick Muller, London: 1980.) Hbk np; pbk £4.50.

PAPAGIANNIS, M. D. (ed.). *Strategies for the Search for Life in the Universe*. Astrophysics and Space Science Library, Vol. 83. Proceedings of Joint Session of the International Astronomical Union held in Montreal, Canada, 1979. Pp.276. Hbk ISBN 90-277-1181-X; pbk ISBN 90-277-1226-3. (Reidel: 1980.) Hbk Dfl. 70, \$37; pbk Dfl. 30, \$14.95.

WYNN-WILLIAMS, C. G. and CRUIKSHANK, D. P. (eds). *Infrared Astronomy*. International Astronomical Union. Symposium No. 96 held in Kona, Hawaii, June 1980. Pp.374. Hbk ISBN 90-277-1227-1; pbk ISBN 90-277-1228-X. (Reidel: 1981.) Hbk Dfl. 90, \$47.50; pbk Dfl. 45, \$23.50.

Physics

BASOV, N. G. (ed.). *Strong and Electromagnetic Interactions of Elementary Particles and Nuclei*. Proceedings of the P. N. Lebedev Physics Institute, Vol. 95. Pp.206. Pbk ISBN 0-306-10965-4. (Consultants Bureau, New York: 1980.) Pbk Np.

CHAKRAVARTY, A. S. *Introduction to the Magnetic Properties of Solids*. Pp.696. ISBN 0-471-07737-2. (Wiley: 1981.) £35, \$81.

ELY, R. V. *Microfocal Radiography*. Pp.295. ISBN 0-12-238140-8. (Academic: 1980.) £28.80.

FELD, M. S. and LETOKHOV, V. S. (eds). *Coherent Nonlinear Optics: Recent Advances*. Topics in Current Physics, Vol. 21. Pp.377. ISBN 3-540-10172-1. (Springer-Verlag: 1980.) DM 83, \$49.

FREYHARDT, H. C. (ed.). *Crystals: Growth, Properties and Applications*. Vol. 4, Organic Crystals, Germanates, Semi-Conductors. Pp.221. ISBN 3-540-10298-1. (Springer-Verlag: 1980.)

HAMMOND, P. *Energy Methods in Electromagnetism*. Pp.180. ISBN 0-19-859328-7. (Clarendon Press/Oxford University Press: 1981.) £18.

HELLWEGE, K.-H. and HELLWEDGE, A. M. (eds). *Landolt-Börnstein. Numerical Data and Functional Relationships in Science and Technology*. Group III: Crystal and Solid State Physics. Vol. 7, Crystal Structure Data of Inorganic Compounds. Pp.330. ISBN 3-540-10147-0. (Springer-Verlag: 1980.) DM 425, \$250.80.

HUNING and HUNG-YUAN. *Proceedings of the 1980 Guanzhou Conference on Theoretical Particle Physics*, Vols. 1 & 2. Pp.1765. ISBN 0-442-20273-3. (Science Press, Beijing, China/Van Nostrand Reinhold: 1980.) Two vols at \$89.50.

KAMINSKII, A. A. *Laser Crystals*. Their Physics and Properties. Springer Series in Optical Sciences, Vol. 14. Pp.456. ISBN 3-540-09576-4. (Springer-Verlag: 1981.) DM 128, \$75.60.

KAO, D. C. and HWANG, W. *Electrical Transport in Solids with Particular Reference to Organic Semiconductors*. International Series in the Science of the Solid State, Vol. 14. Pp.663. ISBN 0-08-023973. (Pergamon: 1981.) £50, \$120.

KONDRATIEV, V. N. and NIKITIN, E. E. *Gas-Phase Reactions*. Kinetics and Mechanisms. Pp.241. ISBN 3-540-09956-5. (Springer-Verlag: 1981.) DM 118, \$69.70.

LE GRAND, Y. and EL HAGE, S. G. *Physiological Optics*. Springer Series in Optical Sciences, Vol. 13. Pp.338. ISBN 3-540-099190. (Springer-Verlag: 1980.) DM 78, \$46.10.

LÉVY, M. *et al.* (eds). *Quarks and Leptons Cargèse 1979*. NATO Advanced Study Institutes Series B. Physics, Vol. 61. Pp.720. ISBN 0-306-40560-1. (Plenum: 1981.) \$75.

LIBBY, P. A. and WILLIAMS, F. A. (eds). *Turbulent Reacting Flows*. Topics in Applied Physics, Vol. 44. Pp.243. ISBN 3-540-10192-6. (Springer-Verlag: 1980.) DM 84, \$49.60.

PETIT, R. (ed.). *Electromagnetic Theory of Gratings*. Topics in Current Physics, Vol. 22. Pp.284. ISBN 3-540-10193-4. (Springer-Verlag: 1980.) DM 65, \$38.40.

RICKAYZEN, G. *Green's Functions and Condensed Matter*. Techniques of Physics, Vol. 5. Pp.357. ISBN 0-12-587950-4. (Academic: 1980.) £22.80, \$52.50.

RING, P. and SCHUCK, P. *The Nuclear Many-Body Problem*. Pp.716. ISBN 3-540-09820-8. (Springer-Verlag: 1980.) DM 86, \$50.80.

TANNER, B. K. and BOWEN, D. K. (eds). *Characterization of Crystal Growth Defects by X-Ray Methods*. NATO Advanced Study Institutes Series. Series B Physics, Vol. 63. Proceedings of the NATO Advanced Study Institute held August 1989, at Durham University U.K. Pp.589. ISBN 0-306-40628-4. (Plenum: 1981.) \$65.

TIRAPUGUI, E. (ed.). *Field Theory, Quantization and Statistical Physics*. In Memory of Bernard Juvet, Vol. 6. Pp.348. ISBN 90-277-1128-3. (Reidel: 1980.) Dfl. 90, \$47.50.

YAROSLAVSKII, L. R. and MERZLYAKOV, N. S. *Methods of Digital Holography*. Pp.171. ISBN 0-306-10963-8. (Plenum: 1980.) \$45.

Chemistry

BAGOTZKY, V. S. and SKUNDIN, A. M. *Chemical Power Sources*. Pp.382. ISBN 0-12-072650-5. (Academic: 1980.) £26.80, \$64.50.

BOSHKE, F. L. (ed.). *Topics in Current Chemistry*. Vol. 93, Van der Waals Systems. Pp.136. ISBN 3-540-10058-X. (Springer-Verlag: 1980.) DM 68, \$40.20.

BROWN, S. S. and SUNDERMAN, F. W. JR. (eds). *Nickel Toxicology*. Proceedings of the 2nd International Conference September 1980, Swansea, Wales, Pp.193. ISBN 0-12-137680-X. (Academic: 1980.) \$36.

Biological Sciences

BERTON, A. D. and ODOM, D. B. *Nuclear Medicine: A Comprehensive Bibliography*. Biomedical Information Guides, Vol.2. Pp.349. ISBN 0-306-65178-5. (IFI/Plenum Data Company: 1980.) \$85.

BERTON, A. D. and ODOM, K. B. (eds). *Smoking and Health: A Comprehensive Bibliography*. Biomedical Information Guides, Vol. 3. Pp.530. ISBN 0-306-65184-X (IFI/Plenum Data Company: 1980.) \$95.

BINDMAN, L. and LIPPOLD, O. *The Neurophysiology of the Cerebral Cortex*. Pp.495. ISBN 0-7131-4360-6. (Edward Arnold: 1981.) £47.50.

BLOWER, J. G., COOK, L. M. and BISHOP, J. A. *Estimating the Size of Animal Populations*. Pp.128. Hbk ISBN 0-04-591017-0; pbk ISBN 0-04-591018-9. (Allen & Urwin, London: 1981.) Hbk np; pbk £4.95.

CALLAHAN, P. S. *The Soul of the Ghost Moth*. Paths of a Naturalist. Pp.104. ISBN 0-8159-6309-2. (The Devin-Adair Company, Old Greenwich, Connecticut: 1981.) \$7.95.

CASE, D. B., SONNENBLICK, E. H. and LARAGH, J. H. (eds). *Captopril and Hypertension*. Pp.236. ISBN 0-306-40532-6. (Plenum: 1980.) Np.

CHIVERS, D. J. (ed.). *Malayan Forest Primates: Ten Years' Study in Tropical Rain Forest*. Pp.364. ISBN 0-306-40626-8. (Plenum: 1980.) \$42.50.

COBB, J. S. and PHILLIPS, B. F. (eds). *The Biology and Management of Lobsters*. Vol. 1, Physiology and Behaviour. Pp.463. ISBN 0-12-177401-5. (Academic: 1980.) \$55.

CULLEN, J. *Notes from the Royal Botanic Garden, Edinburgh*. Vol. 39, No. 1, Revision of *Rhododendron* 1. Pp.207. Pbk ISBN 0-11-491649-7. (HMSO, Edinburgh: 1980.) Pbk £7.50.

CUMMING, G. and BONSIGNORE, G. (eds). *Pulmonary Circulation in Health and Disease*. Proceedings of the Symposium held at the Ettore Majorana Centre for Scientific Culture, Erice, Sicily, Italy, July 1979. Pp.442. ISBN 0-306-40473-7. (Plenum: 1980.) Np.

DAWES, E. A. *Quantitative Problems in Biochemistry*. 6th Edn. Pp.335. ISBN 0-582-44402-0 (Longman: 1980.) £12.95.

De BRABANDER, M. and De MEY, J. (eds). *Microtubules and Microtubule Inhibitors 1980*. Janssen Research Foundation Series Vol. 3. Proceedings of the 2nd International Symposium, Beerse, Belgium, August 1980. Pp.570. ISBN 0-444-80305-X. (Elsevier/North-Holland Biomedical: 1980.) \$79, Dfl. 162.

ENNA, S. J. and YAMAMURA, H. I. (eds). *Neuro-transmitter Receptors*. Part 1, Amino Acids, Peptides and Benzodiazepines. Receptors and Recognition. Series b, Vol. 9. Pp.212. ISBN 0-412-16250-4. (Chapman & Hall: 1980.) £15.

FOSTER, R. W. (ed.). *1200 Multiple Choice Questions in Pharmacology*. Pp.176. Flexi ISBN 0-407-00192-1. (Butterworths: 1980.) £3.50.

FOSTER, R. W., COX, B. *et al.* (eds). *Basic Pharmacology*. Pp.348. Flexi ISBN 0-407-00170-0. (Butterworths: 1980.) £6.95.

FOUGEREAU, M. and DAUSSET, J. (eds). *4th International Congress of Immunology*. Immunology 80. Progress in Immunology IV. Pp.1280. Pbk ISBN 0-12-262940-X. (Academic: 1980.) Pbk £48.

FRANKEL-CONRAT, H. and WAGNER, R. R. (eds). *Comprehensive Virology*. Vol. 16, Virus-Host Interactions: Viral Invasion, Persistence, and Disease. Pp.372. ISBN 0-306-40488-5. (Plenum: 1980.) Np.

GLYNN, L. E. and STEWARD, M. W. (eds). *Antibody Production*. Pp.231. Flexi ISBN 0-471-27916-1. (Wiley: 1980.) £4.50, \$13.

GLYNN, L. E. and STEWARD, M. W. (eds). *Structure and Function of Antibodies*. Pp.306. Flexi ISBN 0-471-27917-X. (Wiley: 1981.) £4.50, \$13.

GOLDBERGER, R. F. (ed.). *Biological Regulation and Development*. Vol. 2, Molecular Organisation and Cell Function. Pp.620. ISBN 0-306-40486-9. (Plenum: 1980.) \$25.

GOODWIN, R. W. *The Biochemistry of the Carotenoids*. Vol. 1, Plants. 2nd Edn. Pp.377. ISBN 0-412-21690-6. (Chapman & Hall: 1980.) £27.50.

GREEN, E. L. *Genetics and Probability in Animal Breeding Experiments*. A Primer and Reference Book on Probability, Segregation, Assortment, Linkage and Mating Systems for Biomedical Scientists Who Breed and Use Genetically Defined Laboratory Animals for Research. Pp.271. ISBN 0-333-27243-9. (Macmillan Press, London and New York: 1981.) £20.

HERVEY, G. F. and HEMS, J. *The Goldfish*. Pp.271. Flexi ISBN 0-571-11611-6. (Faber & Faber, London: 1981.) £3.25.

• HILLMAN, H. and SARTORY, P. The Living Cell. A Re-examination of its Fine Structure. Pp.112. Hbk ISBN 0-906527-02-3; pbk ISBN 0-906527-01-5. (Packard Publishing, Chichester, Sussex: 1980.) Hbk £6.95; pbk £2.95.

HOFFBRAND, A. V. and PETTIT, J. E. Essential Haematology. Pp.236. Flexi ISBN 0-632-00679-X. (Blackwell Scientific: 1980.) £7.50.

HUDSON, L. and HAY, F. C. Practical Immunology. 2nd Edn. Pp.359. Flexi ISBN 0-632-00353-6. (Blackwell Scientific: 1980.) £10.

INGRAM, D. S. and HELGESON, J. P. (eds). Tissue Culture Methods for Plant Pathologists. Pp.272. ISBN 0-632-00715-X. (Blackwell Scientific: 1980.) £15.50.

JOINER, J. N. (ed.). Foliage Plant Production. Pp.614. ISBN 0-13-322867-3. (Prentice-Hall: 1981.) Np.

KOEPCHEN, H. P., HILTON, S. M. and TRZEBSKI, A. (eds). Central Interaction Between Respiratory and Cardiovascular Control Systems. Pp.244. Flexi ISBN 3-540-09948-4. (Springer-Verlag: 1980.) DM 46, £27.20.

KRINKE, G. and HESS, R. Histopathology of Experimental Toxic Neuropathies. Lectures in Toxicology, 1. Pp.6. + slides. ISBN 0-08-025703-8. (Pergamon: 1980.) £25, \$60.

LANE, D. R. (ed.). Jones's Animal Nursing. 3rd Edn. Hbk ISBN 0-08-024945-0; flexi ISBN 0-08-24944-2. (Pergamon: 1980.) Hbk £30, \$67.50; flexi £16, \$36.

LIND, E. M. and BROOK, A. J. A Key to the Commoner Desmids of the English Lake District. Pp.123. Flexi ISBN 0-900386-40-1. (Freshwater Biological Association, Ambleside, Cumbria: 1980.) Np.

MacGREGOR, E. A. and GREENWOOD, C. T. Polymers in Nature. Pp.391. ISBN 0-471-27762-2. (Wiley: 1980.) £19.50, \$58.50.

MANDAL, A. K. and BOHMAN, S.-O. (eds). The Renal Papilla and Hypertension. Pp.237. ISBN 0-306-40506-7. (Plenum: 1980.) Np.

MILLER, A. G. *et al.* Notes from the Royal Botanic Garden, Edinburgh. Vol. 38, No.3. Pp.541. Pbk ISBN 0-11-491663-2. (HMSO, Edinburgh: 1980.) £19.

OSAWA, S. *et al.* Genetics and Evolution of RNA Polymerase, RNA and Ribosomes. Proceedings of the Oji International Seminar, Hokkaido, August — September 1979. Pp.666. ISBN 0-444-80288-6. (Elsevier/North-Holland Biomedical: 1981.) \$115, Dfl. 236.

PEXIEDER, T. (ed.). Mechanisms of Cardiac Morphogenesis and Teratogenesis. Perspectives in Cardiovascular Research, Vol. 5. Pp.528. ISBN 0-89004-460-0. (Raven: 1980.) \$48.

PFAFF, D. W. Estrogens and Brain Function: Neural Analysis of a Hormone-Controlled Mammalian Reproductive Behaviour. Pp.281. ISBN 3-540-90487-5. (Springer-Verlag: 1980.) Np.

POND, W. G. *et al.* (eds). Animal Agriculture. Research to Meet Human Needs in the 21st Century. Pp.355. ISBN 0-86531-032-7. (Westview, Colorado: 1980.) \$20.

ROBERTS, A. and BUSH, B. M. H. (eds). Neurones Without Impulses: Their Significance for Vertebrate and Invertebrate Nervous Systems. Pp.290. Hbk ISBN 0-521-23364-X; pbk ISBN 0-521-29935-7. (Cambridge University Press: 1981.) Hbk £25; pbk £10.95.

ROBERTSON, A. (ed.). Selection Experiments in Laboratory and Domestic Animals. The Proceedings of a Symposium. Commonwealth Agricultural Bureaux. Pp.245. Flexi ISBN 0-85198-461-4. (Commonwealth Bureau of Animal Breeding and Genetics, Edinburgh, Scotland.)

ROGERS, H. J., PERKINS, H. R. and WARD, J. B. Microbial Cell Walls and Membranes. Pp.564. ISBN 0-412-12030-5. (Chapman & Hall: 1980.) £30.

SEGAL, L. A. (ed.). Mathematical Models in Molecular and Cellular Biology. Pp.757. ISBN 0-521-22925-1. (Cambridge University Press: 1981.) £45.

SIMIC, M. G. and DAREL, M. (eds). Autoxidation in Food and Biological Systems. Proceedings of the Workshop held at the U.S. Army Natick Research and Development Command, Natick, Massachusetts, October 1979. Pp.659. ISBN 0-306-40561-X. (Plenum: 1980.) Np.

SPILLANE, J. D. The Doctrine of the Nerves. Chapters in the History of Neurology. Pp.467. ISBN 0-19-261135-6. (Oxford University Press: 1981.) £25.

SPILLER, G. A. and KAY, R. M. (eds). Medical Aspects of Dietary Fiber. Pp.299. ISBN 0-306-40507-5. (Plenum: 1980.) Np.

STEINBERG, A. G. and COOK, C. E. The Distribution of the Human Immunoglobulin Allotypes. Pp.250. ISBN 0-19-261181-X. (Oxford University Press: 1981.) £27.50.

TAN, K. *et al.* Notes from the Royal Botanic Garden, Edinburgh. Vol. 38, No. 2. Pp.371. Pbk ISBN 0-11-491662-4. (HMSO, Edinburgh: 1980.) Pbk £21.

TANABE, Y., TANAKA, K. and OOKAWA, T. (eds). Biological Rhythms in Birds: Neural and Endocrine Aspects. Pp.373. ISBN 3-540-10311-2. (Springer-Verlag: 1980.) DM 86, \$50.60.

VARGA, A. C. The Main Issues in Bioethics. Pp.225. Flexi ISBN 0-8091-2327-4. (Paulist Press, Ramsey, New Jersey: 1980.) \$8.95.

Applied Biological Sciences

COLLIER, L. H. and OXFORD, J. (eds). Developments in Antiviral Therapy. Based on the Proceedings of a Symposium held at the London Hospital Medical College, 30 November 1979. Pp.291. ISBN 0-12-181150-6. (Academic: 1980.) £18.60, \$43.

COSTA, E., DI CHIARA, G. and GESSA, G. L. GABA and Benzodiazepine Receptors. Advances in Biochemical Psychopharmacology, Vol. 24. Pp.321. ISBN 0-89004-530-5. (Raven: 1980.) \$36.

KREIER, J. P. (ed.). Malaria. Vol. 2, Pathology, Vector Studies, and Culture. Pp.328. ISBN 0-12-426102-7. (Academic: 1980.) \$38.50.

LAWSON, A. M., LIM, C. K. and RICHMOND, W. (eds). Clinical Developments in the Clinical Applications of HPLC, GC and MS. Proceedings of a Symposium held at the Clinical Research Centre, Harrow, Middlesex, May/June 1979. Pp.301. ISBN 0-12-439650-X. (Academic: 1980.) £21.40, \$49.50.

MOSTOFI, F. K., SESTERHENN, I. and SOBIN, L. H. (eds). Histological Typing of Prostate Tumours. International Histological Classification of Tumours, No. 22. Pp.25 + 64 colour photomicrographs. ISBN 92-4-176021-4. (WHO, Geneva: 1980.) Sw fr. 109.

SERROU, B. and ROSENFELD, C. (eds). International Symposium on New Trends in Human Immunology and Cancer Immunotherapy. Pp.1073. Flexi ISBN 2-7040-0376-9. (Doin Éditions, Paris: 1980.) Np.

Psychology

CUSHMAN, D. P. and McPHEE, R. D. (eds). Message-Attitude-Behavior Relationship. Theory, Methodology and Application. Pp.339. ISBN 0-12-199760-X. (Academic: 1980.) \$27.

WALLSTEN, T. S. Cognitive Processes in Choice and Decision Behavior. Pp.285. ISBN 0-8959-054-X. (Lawrence Erlbaum, New Jersey: 1980.) Np.

WEISBERH, R. W. Memory, Thought, and Behavior. Pp.458. Flexi ISBN 0-19-502583-0. (Oxford University Press: 1980.) £8.95.

Sociology

AMANN, A. (ed.). Open Care for the Elderly in Seven European Countries: A Pilot Study in the Possibilities and Limits of Care. Pp.225. ISBN 0-08-025215-X. (Pergamon: 1980.) \$33.75, £15.

ILLSLEY, R. Professional or Public Health? Sociology in Health and Medicine. Pp.184. ISBN 0-90057433-X. (Nuffield Provincial Hospitals Trust, London: 1980.) £6.

LINDSEY, D. (ed.). Preventive Approaches to Child Welfare. A Special Issue of Children and Youth Services Review. Pp.476. Flexi ISBN 0-08-026108-6. (Pergamon: 1980.) £6.

SMITH, P. K. and CONNOLLY, K. J. The Ecology of Preschool Behaviour. Pp.383. ISBN 0-521-22331-8. (Cambridge University Press: 1981.) £25.

History of Science

BUCCIARELLI, L. L. and DWORSKY, N. Sophie Germain. An Essay in the History of the Theory of Elasticity: Studies in the History of Modern Science, Vol. 6. Pp.143. Hbk ISBN 90-277-1134-8; pbk ISBN 20-277-1135-6. (Reidel: 1980.) Hbk Dfl. 60, \$31.50; pbk Dfl. 30, \$15.75.

WHEELER, J. A. Albert Einstein: His Strength and His Struggle. The Twentieth Selig Brodetsky Memorial Lecture. Pp.14. (Leeds University Press: 1980.) Pbk £10.95.

General

ARMSTRONG, E. S. Oracles. A Design System for Linear Multivariable Control. Control and Systems Theory. A Series of Monographs and Textbooks, Vol. 10. Pp.256. ISBN 0-8247-1239-0. (Dekker: 1980.) SwFr. 75.

ASQUITH, P. D. and GIERE, R. N. PSA 1980. Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 1. Pp.370. Hbk ISBN 0-917586-14-X; pbk ISBN 0-917586-13-1. (Philosophy of Science Association, Michigan State University, Michigan: 1980.) Hbk \$9.50; pbk \$7.50.

BALLENTYNE, D. W. G. and LOVETT, D. R. A Dictionary of Names Effects and Laws in Chemistry, Physics and Mathematics. 4th Edn. Pp.346. Hbk ISBN 0-412-22380-5; pbk ISBN 0-412-22390-2. (Chapman & Hall: 1980.) Hbk £15; pbk £7.95.

BARNABY, F. Prospects for Peace. Pp.88. Flexi ISBN 0-08-027398-X; hbk ISBN 0-08-027399-8. (Pergamon: 1980.) Flexi \$11.50; hbk \$21.50.

BENNETT, D. *et al.* The Harvey Lectures. Pp.254. ISBN 0-12-312074-8. (Academic: 1980.) \$21.

BERES, L. R. Apocalypse: Nuclear Catastrophe in World Politics. Pp.315. ISBN 0-226-04360-6. (University of Chicago Press: 1980.) £12.

BERTING, J., MILLS, S. C. and WINTERSBERGER, H. (eds). The Socio-Economic Impact of Microelectronics. International Conference on Socio-Economic Problems and Potentialities of Microelectronics, Zandvoort. Pp.267. ISBN 0-08-026776-9. (Pergamon: 1980.) \$50, £22.25.

BLOWER, C. The Secrets of a Hypnotist in General Practice, Vol. 1. Pp.40. (Hawkesbury Medical Windsor, NSW, Australia: 1980.)

BOGDAN, R. J. (ed.). Keith Lehrer: Profiles. An International Series on Contemporary Philosophers and Logicians, Vol. 2. Pp.255. Hbk ISBN 90-277-1172-0; pbk ISBN 90-277-1173-9. (Reidel: 1980.) Hbk Dfl. 70, \$37; pbk Dfl. 25, \$12.95.

BOLDY, D. (ed.). Operational Research Applied to Health Services. Pp.266. ISBN 0-7099-0380-4. (Croom Helm: 1981.) £15.95.

CAPPER, P. L., CASSIE, W. F. and GEDDES, J. D. Problems in Engineering Soils. 3rd Edn. Pp.287. Flexi ISBN 0-419-11840-3. (E. & F. N. Spon, London: 1980.) £5.75.

CHARALAMBOUS, G. (ed.). The Analysis and Control of Less Desirable Flavours in Foods and Beverages. Pp.358. ISBN 0-12-169065-2. (Academic: 1980.) \$24.50.

CHIBISOVA, O. I. and KOZIAR, L. A. (eds). English-Russian Biological Dictionary. 4th Edn. Pp.732. ISBN 0-08-023163-2. (Plenum: 1980.) £50, \$113.

COLQUHOUN, W. P. and RUTENFRANZ, J. (eds). Studies of Shiftwork. Pp.468. ISBN 0-05066-210-9. (Taylor & Francis: 1980.) £16.50.

CONRAD, J. (ed.). Society, Technology and Risk Assessment. Proceedings of an International Workshop held at Wolfersheim, Germany, June 1979. Pp.303. ISBN 0-12-186450-2. (Academic: 1980.) £15.20, \$35.

DORF, R. C. Modern Control Systems. 3rd Edn. Pp.493. Flexi ISBN 0-201-03147-7. (Addison-Wesley: 1980.) £8.95.

DUNCAN, W. R. (ed.). Soviet Policy in the Third World. Pp.322. ISBN 0-08-025125-0. (Pergamon: 1980.) £17.50.

ELKINGTON, J. The Ecology of Tomorrow's World — Industry's Environment. Pp.311. ISBN 08227-260-X. (Associated Business Press, London: 1980.) £12.

RES PONDENCE

Continued from page 82

Nature Conservation Council has publicly stated that such sites are "gems" (BBC Radio 4, 22 January). He is indeed correct, for, like jewels, they provide neither nourishment nor habitation nor protection (the three necessities of life). But they are worse than jewels for they cannot be traded in time of hardship and instead of representing riches they are a permanent impoverishment to their owners. The usual epithet for such items is "white elephants".

Let us have no more of this folly in the name of science. The value to science of objects or phenomena lies not in themselves but in the information they yield to study. Once this information is recorded and published whatever value remains in the objects or phenomena is of no value to science, for unless the original study was incompetently executed nothing new will be learned by preservation. I do not in the least condone the wilful pursuit of rare organisms, or the careless or selfish pollution of the environment, and I admit the possible future benefits which may be derived by maintaining "gene banks" of rare forms of cultivated organisms, but all that is quite a different matter. Should there be a popular demand for the preservation of areas of wilderness then justice requires that the populace provide compensation to the owners, but let there be no mistaken idea that this cost, which will be measured in hundreds of thousands of pounds annually, is a deserving use of the funds made available to science, nor that compulsion is justified by the supposed advance of knowledge. If any individual or body of scientists presumes that it is only necessary to legislate or to spend enough money in order to halt not only natural selection but also the very evolution of the inanimate world, such presumption is an affront to reason and must bear the condemnation of responsible scientists for misleading those who know no better.

C. MUIR

Department of Zoology,
The University, St Andrews, UK

Origin of cancer

SIR — Cairns's¹ article is a welcome breath of fresh air in trying to put the causes of cancer into a realistic perspective. However, he could have strengthened his case for genetic transposition as a major cause by considering the relevance of DNA double-strand breaks as a molecular mechanism. Over the last few years Leenhouts and Chadwick² have convincingly demonstrated that the biochemical basis of malignancy lies most probably in the generation of double-strand breaks—lesions which cannot be repaired but which may result in genetic transposition. The evidence comes from many sources including those quoted by Cairns as well as radiation damage effects and cell-fusion experiments. Since double-strand breaks are necessary for some mutations, chromosomal abnormalities and genetic transpositions, Occam's razor

would suggest this as a suitable model on which to test Cairns's hypothesis.

D.I. EDWARDS

Department of Paramedical Science,
North East London Polytechnic,
London E15, UK

1. Cairns, J. *Nature* **289**, 353-357 (1981).
2. Leenhouts, M.P. & Chadwick, K.M. *Int. J. Radiat. Biol.* **33**, 357-370 (1978).

SIR — John Cairns's speculative article¹ on the origin of human cancer is open to criticism both for what it says and for what it leaves unsaid.

Garner and Hertzog² have already drawn attention to the tenuous nature of the evidence on which Cairns bases his suggestion that "Chemical mutagens (of the kind detected in the usual tests for mutagenicity) seem unlikely to be rate-limiting components in most human carcinogenesis" and that "large-scale changes in the genome (such as rearrangements and deletions) seem to be more hazardous than the local changes produced by conventional mutagens". They might have put their case even more strongly.

First, the fact that patients with xeroderma pigmentosum (XP) do not show a high incidence of cancer in tissues other than skin may imply only that cancer in these other tissues is not associated with the formation of "UV-like" lesions³ because the essential anomaly in XP seems to be failure to recognize lesions of this kind rather than inability to excise and repair DNA lesions in general.

Second, the association of chromosomal abnormalities with a high incidence of leukaemias, lymphomas and carcinomas in patients with Bloom's syndrome tells us nothing about the aetiology of cancer in the vast majority of the population who do not have Bloom's syndrome. Similarly it seems unjustifiable to assume that most cancer cells have an abnormal karyotype simply because this is often the case with cells from leukaemias, lymphomas, meningiomas and gliomas. Moreover, even if this generalization were true, it is conceivable, as Cairns himself points out, that the observed chromosomal rearrangements could have resulted from "trivial secondary events that occur after all the rate-limiting steps of carcinogenesis have been completed".

Third, it is not clear how the phenomenon of carcinogenesis in experimental animals by implanted films of plastic (and other materials) supports Cairns's argument. The underlying mechanism is still far from clear but, contrary to what Cairns suggests, Karp *et al.*⁴ (whose work he cites) concluded that the hypothesis that tumours develop because the cells have lost the restraining influence provided by contact with each other, was specifically excluded by their observations. It is also worth pointing out that this form of carcinogenesis does not seem to have been reported in man despite the fact that foreign material of various kinds, and sometimes in sheet form, has been used in human surgery for many years (for example, in hernia repair, orthopaedic surgery and vascular surgery).

Finally, even if large-scale changes in the genome were shown to occur in cells from most common human cancers, this would not

exclude the possibility that environmental pollution by chemical mutagens of various kinds, including mutagens detectable by the Ames test, plays an important part in the aetiology of human cancer. Many such mutagens, as Garner and Hertzog point out, and as Cairns admits, can increase the frequency of sister chromatid exchange and cause other chromosomal interactions in mammalian cells. The extent to which such changes are "functionally equivalent to genetic transpositions" seems to be of secondary importance in comparison with the question of whether or not they are associated with a high incidence of cancer.

As regards what is left unsaid, Cairns seems to equate "the creation of a cancer cell" with the development of cancer. Thus, while he recognizes the possible role of DNA repair mechanisms in preventing the emergence of cancer cells, he ignores the possibility, which I have discussed at length elsewhere⁵, that homeostatic mechanisms of another kind may operate to destroy transformed cells or prevent them from dividing, and thus prevent the development of overt cancer. It is clear from the work of Mondale and Heidelberger⁶, and others, which Cairns cites, that there may be a long interval between the application of a mutagen to cells in tissue culture and the emergence on repeated subculture of cells with the hallmarks of malignant transformation. The mechanisms responsible have not been fully elucidated but there seems no reason to doubt that they also operate *in vivo* and play a part in delaying the development of tumours in response to exposure to chemical and physical carcinogens; this, however, does not exclude the possibility that, *in vivo*, other mechanisms are also involved.

The long interval, often of many years' duration, which can occur between the locally successful removal of a primary tumour and the appearance of metastases would seem to imply either that cancer cells can remain dormant, or that cell proliferation may be balanced by cell death, for a long time, and studies of tumour cell population kinetics in both animals and man confirm this conclusion⁵. It has, moreover, been observed that the period of apparent dormancy may be abruptly terminated by experimental⁷ or therapeutic⁸ procedures. If metastatic tumours can be held in check in this way, why not primary tumours? My colleagues and I have recently described a technique which makes it possible to investigate this question directly in an animal model^{9,10}.

MICHAEL WOODRUFF

MRC Clinical and Population
Cytogenetics Unit,
Western General Hospital,
Edinburgh, UK

1. Cairns, J. *Nature* **289**, 353-357 (1981).
2. Garner, R.C. & Hertzog, P.J. *Nature* **289**, 627 (1981).
3. Regan, J.D. & Setlow, R.B. *Cancer Res* **34**, 3318-3325 (1974).
4. Karp, R.D. *et al. J. natn. Cancer Inst.* **51**, 1275-1285 (1973).
5. Woodruff, M.F.A. *The Interaction of Cancer and Host* (Grune & Stratton, New York, 1980).
6. Mondale, S. & Heidelberger, C. *Proc. natn. Acad. Sci. U.S.A.* **65**, 219-225 (1970).
7. Eccles, S.A. & Alexander, P. *Nature* **257**, 52-53 (1975).
8. Woodruff, M.F.A. *Nature* **258**, 776 (1975).
9. Woodruff, M.F.A. *et al. Proc. R. Soc. B* (in the press).
10. I thank Professor H.J. Evans for his helpful comments on this letter.

ATTENTION

Awards

Dr Alvin M. Weinberg (Oak Ridge) and **Sir Rudolf E. Peierls** (Oxford) have been selected to receive the US Department of Energy's Enrico Fermi Award for 1980.

Dr Seymour A. Papert (Massachusetts Institute of Technology) has been awarded the Seventh Marconi International Fellowship for his work in computer education.

The William Bate Hardy Prize of the Cambridge Philosophical Society has been awarded to **Dr C. Milstein** (Fitzwilliam College) for his distinguished work on monoclonal antibody techniques.

The Zoological Society of London has made the following awards for contributions to zoology in 1980: The Scientific Medal to **Dr J. M. Elliot** (Freshwater Biological Association, Ambleside); The Stamford Raffles Award to **Dr A. J. Woakes** (University of Birmingham); The Zoological Society of London Frink Medal for British Zoologists to **Professor W. H. Thorpe** (University of Cambridge); The Prince Philip Prize to **James Burton** (of King's College, Taunton).

The André Lichtwitz Prize (10,000 French Francs) is awarded every year by INSERM for distinguished clinical or fundamental research on calcium and phosphorus metabolism carried out by a French or foreign scientist (or a team). Applicants who have previously applied are allowed to apply again. Details from: Director of INSERM (c/o Mlle C. CHIROL, 101 rue de Tolbiac 75645 Paris Cedex 13, France).

Appointments

Arthur L. Schawlow (Stanford University) has been appointed President of The American Physical Society.

Professor Graham Milbourn has been appointed Director of the National Institute of Agricultural Botany at Cambridge.

Gordon Bell who has been Director of the Royal Observatory, Hong Kong for the past 15 years retired in January, to be succeeded by **Mr J. E. Peacock**. Mr Bell has been appointed to the newly created post of Science Adviser to the Hong Kong Government.

Dr George Stewart, has been appointed to the Chair of Petroleum Engineering at Heriot-Watt University, Edinburgh, in succession to **Professor James Brown**, who retires on 31 March 1981.

The University of Newcastle upon Tyne has made the following appointments: **Dr F. K. Kong** to the Chair of Structural Engineering from 1 September 1981, and **Dr C. J. Hull** to the Chair of Anaesthesia from 1 January 1981 and to the Headship of the Department of Anaesthesia from 1 January 1981 to 30 September 1986.

Dr Anthony G. Atkins has been appointed Professor of Mechanical Engineering at Reading University from 1 March 1981.

Meetings

5-8 October, **Hormones and Cell Regulation**, Bischoffsheim (Dr G. Schultz, Pharmakologisches Institut der Universität Heidelberg, im Neuenheimer Feld 366, D-6900, Heidelberg, Germany).

5 October — 6 November, **Land Use Planning and Environment Applications**, Flagstaff (J. A. Reinemeund, US Geological Survey, National Center — M/S 917, 12201 Sunrise Valley Drive, Reston, Virginia 22092, USA).

6-9 October, **Safety in Communications and Transport**, Genoa (IIC, Via Pertinace, Villa Piaggio, 16125 Genoa, Italy).

5-10 October, **Tropical Ecology and Resources Management**, Shopal (Prof. K. C. Misra, Secretary ISTE, Dept of Botany, Banaras Hindu University, Varanasi-221005, India).

6-10 October, **Microprocessors in Control of Mechanical Function**, Leeds (The Institution of Mechanical Engineers, 1 Birdcage Walk, London SW1, UK).

11-17 October, **Marine Sciences in the Red Sea**, Al-Ghardaqa (Prof. H. A. F. Gohar, 101 Kasr Al-Ainy, St, Cairo, Egypt).

12 October, **Mutagenicity Testing**, London (Mr M. L. Richardson, c/o Mr E. A. Taylor, The Royal Society of Chemistry, Burlington House, London W1, UK).

18-31 October, **Chemical Carcinogenesis**, Erice-Trapani (Dr S. Parodi, Istituto Scientifico per lo Studio e la Cura dei Tumori, Viale Benedetto XV, 10-16132 Genova, Italy).

19-23 October, **Concrete Roads**, London (The Concrete Society, Terminal House, Grosvenor Gardens, London SW1, UK).

26-30 October, **Artificial and Human Intelligence**, Lyon (Dr A. Elithorn, The Royal Free Hospital, Pond St, London SW3, UK).

27-29 October, **Petroanalysis**, London (Miss I. A. McCann, Institute of Petroleum, 61 New Cavendish St, London W1, UK).

GUIDE TO AUTHORS

● Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.

● Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.

● Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.

● 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*.

Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov. (1976).

nature

19 March 1981

US versus science education

The first Reagan budget, well advertised in advance as a device for getting the government off the taxpayer's back and the taxman out of his pocket, has more than lived up to expectations in its mean view of science education, at least as administered by the National Science Foundation. Ever since the heady post-sputnik days, when Washington was thronged with committees giving the Eisenhower Administration advice on how best to meet this unexpected challenge to American technology, the National Science Foundation has been spending a sizeable part of its budget (but still a modest annual sum) on the support of curriculum development in schools and colleges. It can boast of having done more good than foolish works. Zacharias's organization, the Physical Sciences Study Committee, may not have produced the ideal physics curriculum for the average high school, but it was a cheerful demonstration that distinguished academics were willing and able to help with the problems of teaching in the schools. And many young people (not only in the United States) learned a great deal of physics from the textbooks, readers and experiments that the committee devised. In retrospect, however, it may have been even more telling that the experiment proved infectious. Quite soon, in the early 1960s, an army of high-school teachers and a battalion of academics were caught up in the network of committees dedicated to the improvement of science teaching.

Inevitably, there have been excesses and mistakes. If the foundation was less confident in the development of curricula in the social sciences and the humanities than in the physical or biological sciences, it should not shoulder all the blame. Mere money is not sufficient to design a good curriculum. The essential is that good teachers and perceptive academics should be willing, and that school systems should help with administration and facilities. In any case, what matters is not that such and such a set of books should be written as an integrated package (and then, with luck, adopted in the Californian schools) but that teachers should have learned that continual change is necessary but also

possible. The \$13 million the foundation spent on these modest efforts last year will not go far to help the US Treasury to balance its books (or, rather, to live with a deficit of \$41,000 million). Cutting back on this part of the foundation's programme when the high schools are painfully aware of the need for new kinds of teaching seems either unwise or spiteful. Or is the Administration calculating that with Mr Walter Cronkite's new science programme advertised on CBS for the summer, the National Science Foundation can fade gracefully from the scene?

The budget request will also cut the foundation's modest support for students of science and engineering, mostly at the graduate level. The timing is entirely misplaced. Only a few months ago, the Carter Administration was agonizing about the problems of recruiting good people into engineering and especially engineering education. The problem cannot have disappeared since Inauguration Day. And in spite of the largely self-sustained character of American students — the envy of British university administrators (see below) — the United States government is probably still getting value for money from its spending on this good cause. Any Presidential Science Advisor would vouch for that, which may be one of the reasons why President Reagan appears reluctant to appoint such a person.

One of the temptations that afflict incoming administrations everywhere is that of abolishing at least some pre-existing institutions. The National Science Foundation's education programme is a natural candidate for this treatment. What it has been doing is what the market might otherwise have to do. It is also an administrative inconvenience that there should be an organization concerned with education entirely separate from the Department of Education. And for all their clubbiness, people interested in such things as curriculum development do not constitute a constituency of any importance. So why not abolish the institution? The irony, which other administrations have found to their cost, is how often it is necessary to recreate a substitute of some kind.

. . . . and UK versus universities

British universities, after too many months of somnolence, have taken fright. Last week, Sir Alec Merrison, Vice-Chancellor of the University of Bristol but speaking in his capacity as chairman of the Committee of Vice-Chancellors and Principals, accused the government of "a kind of madness". He and a posse of fellow vice-chancellors seem privately to have told the Secretary of State for Education and Science, Mr Mark Carlisle, much the same; by Merrison's account, they made no impression with their complaints that the British government is "profoundly misguided" in its intention to cut state support for the university system by 8.5 per cent in the next three years. Together with the switch to full economic fees for overseas students now under way, state support for the British university system will have fallen by at least 11 per cent, and perhaps by 15 per cent, by the academic year beginning in 1983. Either way, structural changes of some kind will be forced on the system. It is no wonder that people have taken to strong language.

Universities themselves must, however, shoulder some of the responsibility for what is now in prospect. At least since the summer of 1978 (under a different government), it has been clear that structural change was unavoidable. Even the now forgotten promise of "level funding" would not have allowed universities

to keep on recruiting young people to academic posts at a sufficient rate to stay intellectually alive, also competitive. The age distribution of people in tenured posts, skewed by the rapid recruitment in the 1960s of then young academics, ensures that in the 1980s the rate of retirement among British academics (now one per cent a year) will be less than a third of that needed to permit recruitment at a rate corresponding to a steady state. British universities have talked at length about the problem; none has found a solution. The government's decision, two years ago, that universities must collect larger tuition fees from overseas than from home students, whose first effects are becoming apparent only in this academic year, has had the effect of distributing penury at random. Again, most universities have put on hair shirts to wait for whatever the fates might bring. The outstanding exception is the London School of Economics, which has actually increased the number of its new overseas customers by the simple expedient of telling prospective students (and their governments) about its own virtues. The new cut goes deeper. The government intends simply to take cash out of the system, and there is no alternative source in sight. The damage that will be done would have been more easily contained if the university system had responded much more imaginatively to the previous

crises. There has been plenty of warning.

The government's behaviour is even less creditable, but consistent with its brief record. Overseas students' fees were increased by fiat, without consultation (and without much thought). The result is not merely the arbitrariness with which different universities are affected but the patchwork of petty anomalies and injustices with which university administrators have to struggle. The latest cut has been deviously introduced. Last December, when the government casually announced a cut of 3.5 per cent in next year's budget, nobody mentioned that there was worse to come. And last week the announcement of two further instalments (in 1982 and 1983) was tucked away in the Treasury's annual forecast of government expenditure. Was somebody hoping that nobody would notice? Or that the issue would be subsumed in the generalized sense of grievance which this year's higher taxes have evoked? Whatever the truth, there seems to have been only the most cursory consultation with the universities about the consequences of the new cuts and no serious attempt to discover the least damaging ways of saving money. In short, there is no explicit policy for the universities. On all the evidence, there seems not even to be a secret policy. Ministries involved with higher education utter threats and then, several weeks later, deny them with reassuring promises. A few weeks ago, for example, the government had given the universities to understand that they should not try to solve their money problems by taking in extra students; last week's documents on spending promises that they can take as many as they can teach. Is it not time that Parliament forced the government to say what is in its mind? Have the universities no friends left at Westminster?

Ironically, the academic community is probably as keenly aware of the government's economic problems as any other in British society, with the possible exception of the bankers, themselves afflicted by a retrospective capital levy in this year's British budget. The government's central objective is to borrow less from the commercial markets, to which end spending less is important. The new cuts will yield £80 million a year and the increased fees for overseas students nearly as much. However the "saving" is shared out among the British universities, the consequences will be bad for the quality of British university education. The government's expenditure plan says laconically (in a one-sentence paragraph) that "the plans assume a significant tightening of staffing standards", which is a euphemism for a further decline in the ratio of teachers to students. The plan also says that the universities collectively will admit fewer students in the years ahead. The plain implication is that the system will have to shed teachers. That, in the last resort, is why the vice-chancellors are up in arms.

So how should the universities respond? As they should have done in the previous crises of the past three years, by organizing themselves for flexibility. With a few exceptions, the objectives of most university institutions in Britain are similar — the pursuit of the highest academic standards in the teaching of undergraduates and of excellence in research. For most universities and almost all academics, the ideal is Oxbridge. And there is very little encouragement, in the way the government subvention is disbursed on the recommendations of the University Grants Committee, for them to aim at distinctive goals. The result is that too many British universities are trying to do the same things with such inappropriate resources that they are bound to fail.

Would it not, however, be preferable that they should aim at excellence in what they already do well? The system includes half a dozen institutions which are what would be called liberal arts colleges in the United States but which nevertheless waste endless time and money on graduate education in the traditional pattern without recognizing that their graduate schools will always be too small to be viable. It would be better for the institutions concerned, their students and the system as a whole, if this were recognized. The strong civic universities in Britain, foundations of the late nineteenth century, are another distinctive category of universities. Not so long ago, they drew a great deal of strength (and substantial amounts of cash) from the cities in which they were sited. Although local loyalties are still in many places strong,

they are nowhere as strong as they used to be when it was no shame for a young person to become a student at the nearest university. In the serious crisis about to inflict itself on them, the civic universities will regret the way that the old loyalties have been eroded. Some of them may even find themselves in sneaking sympathy with the plan which the government is said to be hatching for replacing the present system of maintenance grants for students (over and above tuition fees) with a smaller grant and the offer of a loan. There could hardly be a better device for channelling students towards the nearest university, while stronger local links would give these universities opportunities to increase their present interest in the retraining of local adults — professionals, technicians, even craftsmen.

The benefits of reorganization for diversity are several, not least the direct economies there would be once the universities stopped aping each other. In such a pattern, the system could hope to match courses to students more expertly than at present, while a clear knowledge of their distinctive roles would make it possible for individual institutions to tell a more convincing tale when asked (as they will be frequently in the years ahead) why they continue to exist. But how is such a pattern (or any other pattern) to be brought about? The conventional wisdom among academics is that universities are so incapable of making long-term policy, and of imposing change upon themselves, that change must be forced from outside, either by the University Grants Committee or by some seemingly impartial committee made up of academics from elsewhere. This is too depressed a view.

Although the occasions in the past when individual universities have changed radically have been times of plenty, there is no reason why the penury now in prospect should not also be a spur to change. Sheer self-interest dictates that it should be, but the University Grants Committee could do much to help. Institutions willing to take painful steps — closing, say, a particular department, or all graduate schools — need and deserve compensation. The snag, in present circumstances, is that an institution making substantial savings in its internal operations cannot be sure that it will benefit in the long run. The University Grants Committee can always in its wisdom recommend that the funds should be spent elsewhere in the system. So would it not make sense, for the duration of whatever transitional period lies ahead, that institutions making voluntary sacrifices should benefit from some kind of salve for their wounds? The University Grants Committee, more used to sharing resources according to its own hidden definitions of equity, is not used to the encouragement of entrepreneurship. But even the University Grants Committee will have to change before this crisis is over.

So too will many other institutions. Even if the outcome of the next few years is a concentration of graduate education in a subset of the universities, it would be intolerable if academics and graduating students from elsewhere were denied opportunities for research. This would require that sources of research support, the research councils in particular, should change. Then, given the need for compatibility within a mobile and diversified system, most universities would find it necessary to restructure their undergraduate curricula. Many, no doubt, would rediscover the virtues of the four-year undergraduate courses, of less specialized (if still competitive) entry requirements and of less narrow courses of study. In other words, there may be benefits in the present adversity.

In the wake of last week's shock to British universities, many dispondent academics will consider this cheerful outcome a delusion. Is it not more likely, they will ask, that the University Grants Committee will recommend that misery should be evenly distributed, with the result that all universities become less good? That, certainly, is possible. It is, however, unlikely that all universities within the system will be equally passive. When they have recovered their breath, the strongest among the universities will be thinking of how to secure for themselves a greater measure of independence than they have enjoyed for years. In due course, all those that survive will follow suit. For what the latest crisis in the British system shows is that universities, jealous of their claims to autonomy, are safest when they are also independent.

Kiss of life for Clinch River breeder

Reagan ditches Carter policy on proliferation

Washington

Despite the objections of his economic advisers, President Ronald Reagan has decided to pay off some of his political debts by letting construction work begin on the liquid metal fast-breeder reactor at Clinch River in Tennessee, which critics claim is both economically unnecessary and technically out of date. However, Mr Reagan is holding back from providing further federal support for the reprocessing of spent nuclear fuel until it becomes clear whether the private sector is prepared to take on this responsibility.

Both the fast breeder programme and commercial spent-fuel reprocessing had been held in check by President Carter on the grounds that they contributed to the growth of the "plutonium economy" and thus raised the risk of the proliferation of nuclear weapons. But both decisions have come under strong attack from the nuclear industry as an unnecessary restriction on the growth of nuclear power, an attitude known to be held by the new Energy Secretary.

In his budget request to Congress, the full details of which were released last week, President Reagan is requesting a total of \$737 million for support of fast-breeder research in the fiscal year 1982. A substantial proportion of this increase will go towards the start of construction work on the Clinch River reactor as a "demonstration project". Development funds have already paid for the design and much of the component fabrication, but construction had previously been held up by President Carter's request to the Nuclear Regulatory Commission not to issue a construction licence.

The Clinch River reactor has a high symbolic value for the US nuclear industry. Many used it to focus their opposition to Mr Carter's general non-proliferation policies. Others argued against the Carter Administration's conclusion that the declining rate of increase in the demand for electricity, and a short-term surplus in domestic uranium supplies, reduced the need for an intensive breeder development programme.

Environmentalists now point out that they have a new — and perhaps surprising — ally, Mr David Stockman, director of the Office of Management and Budget and the chief architect of Mr Reagan's austerity budget. Using the same arguments put forward by the Carter Administration, Mr Stockman circulated a letter to fellow con-

gressmen in 1977 describing plans to subsidize the construction of the Clinch River reactor as "totally incompatible with the free market approach to energy policy".

Critics of fast-breeder reactors, such as the Natural Resources Defense Council, claim that in order to find the money for increased spending on nuclear programmes in general, including the liquid metal fast-breeder reactor, the Department of Energy is intending to cut back significantly in research and development on alternative energy sources. It is proposed to cut the solar energy research budget, for example, from \$625 million in 1981 to \$241 million in 1982, and research and development in conservation from \$558 to \$195 million, while research on fission reactors would increase from \$1,166 to \$1,247 million.

Even the Office of Management and Budget is said to have argued that there is

an inconsistency in cutting off support for "demonstration" projects in solar and fossil energy, while supporting just such an approach for fast breeders; Mr Stockman is said to have put up a fierce battle against Clinch River funding, although he was less opposed to the increased budget request for design work on a larger, more advanced fast breeder.

In a report published by the subcommittee last week, a former member who left Congress last autumn, Republican congressman John Wydler of New York, reported on a visit to the British fast breeder programme, and used the British (and French) experience to bolster his argument that completion of Clinch River is the logical next step for the United States.

Mr Wydler also reported that British scientists were attracted to the idea of cooperation with the United States, a move

British academics up in arms

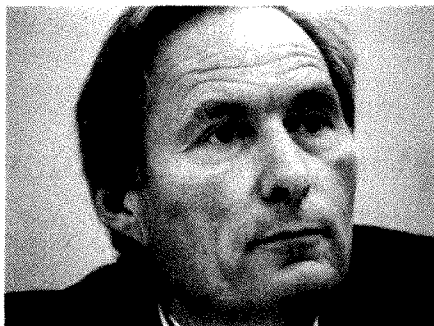
British universities were in a state of shock last week after the government's announcement of its spending plans for the next three years. Spending on home students is to be cut by 8 per cent in real terms by 1983–84. The Committee of Vice-Chancellors and Principals estimates that the total reduction in university income could be as high as 15 per cent when the shortfall in overseas students, expected after the government's decision to charge them economic fees, is taken into account. Sir Alec Merrison, chairman of the Committee of Vice-Chancellors, last week described the expenditure plan as madness. He said he was not aware that the government was reducing the universities' grant for any reason other than saving public money.

The cuts will undoubtedly force the universities to review the services they can offer, which is a fact acknowledged by the government in its white paper on expenditure published on budget day, 10 March. The vice-chancellors and the University Grants Committee seem to have persuaded the Secretary of State for Education that money should be set aside for the scaling down operation. But, instead of making more money available, as the universities must have hoped, the government announced on 13 March that £20 million should be taken out of the 1981–82 recurrent grant — at £979 million already 3 per cent less than for 1980–81 — for the purpose.

The Committee of Vice-Chancellors estimates that making necessary economies could be more expensive than maintaining the universities at their present level. The problem is that after several years of static budgets, further economies cannot be made without reducing staff. Salaries and related costs at present consume 72 per cent of the recurrent grant, more than half of

that going on academic pay. Redundancies, which would break the contracts of employment of tenured academics (approximately 80 per cent hold tenure), would involve the universities in costly litigation and large redundancy payments. And the committee's estimate that one in eight jobs will have to go would certainly mean the closure of departments and possibly entire universities. Nobody yet knows whether the £20 million set aside will be enough to help meet these costs.

The University Grants Committee will begin dividing what money there is between institutions in May. It is not thought to



"Madness" says Merrison

favour cutting any university's grant completely, but neither does it want to spread the money too thinly. Possible alternatives are to close selected departments — a new move for the committee, which normally leaves individual universities to run their own affairs — or to restructure the system, leaving some universities to develop as centres of excellence and turning others into something akin to American liberal arts colleges. The trouble is that the committee has never had to be so *dirigiste* in the past and has not formed a proper mechanism to take the necessary decisions.

Judy Redfearn

which had previously been endorsed by the committee, particularly because of problems with leakage in the existing British fast reactor steam generator components. Mr Wydler also said, however, that UK Atomic Energy Authority officials felt that cooperative agreements with the United States presented unique difficulties because of the Freedom of Information Act and the Atomic Energy Act requiring the disclosure of proprietary business information.

In contrast with his defeat in the fast breeder debate, Mr Stockman seems to have won on the other point of dispute with nuclear advocates in the Department of Energy, namely renewed federal support for spent fuel reprocessing from commercial reactors. The nuclear industry has been pushing for the reopening of the Barnwell reprocessing plant in South Carolina, which was denied funds by the Carter Administration; no money for commercial reprocessing is, however, included in the budget request, although there are funds for the West Valley Demonstration Project.

Backing up its general stance in favour of reprocessing, however, the Reagan Administration has decided to discontinue the spent-fuel programme intended to provide storage facilities away from power station sites. This programme is being refocused within the Department of Energy to concentrate on the development of alternative spent-fuel storage technologies which can be used by utilities at existing reactor sites.

Although the Administration has set its face against subsidies for commercial reprocessing, it has in the past few days become known that plans to dismantle at least this aspect of the Carter anti-proliferation policy are being worked out.

David Dickson

German big science

Win some, lose some

Germany should spend an extra DM 800 million (£170 million) on big science projects over the next ten years, recommends the committee set up a year ago by the previous minister of science, Volker Hauff — and another DM 425 million of projects should be accounted for under the current budget.

The "big science" committee, chaired by Professor Klaus Pinkau (who was recently appointed director of the Max Planck Institut für Plasmaphysik at Garching) reported last week to Minister Andreas von Bülow on a shopping list of ten major projects (*Nature* 3 July 1980, p.8) ranging from high-energy physics to geology, which, if they were all undertaken now, would have cost Germany at least DM 2,015 million (430 million). On being given the task, Pinkau said it would be a "work of art" to steer the right course between scientific worthiness and the government's

ability to pay — for the committee was given no financial guidelines. In the interim, the German economy has weakened, and the Pinkau committee felt obliged to take cost strongly into account.

However, some sleight of hand reduces the bill. The committee recommends that Germany support LEP, the large electron-positron collider which physicists would like to see as the next major facility at CERN, the European subnuclear physics laboratory in Geneva. LEP will cost Germany DM 350 million; but as it is likely to be built within CERN's current budget and staff, the committee reckons it will mean no extra expenditure. Similarly, a DM 75 million replacement for Meteor, Germany's principal and ageing oceanographic research vessel, can be paid for within the current geosciences budget.

Of the other projects, the committee says the refitting and expansion of the BER II research reactor at the Hahn-Meitner Institut, Berlin, costing DM 47 million, should be undertaken immediately; so should the construction of the DM 33 million SUSE, a 250 MeV superconducting cyclotron for heavy ion physics proposed by the Technical University of Munich — provided the university can raise part-finance from the regional government of Bavaria, and make the accelerator a national facility. An equivalent facility proposed by the Jülich nuclear laboratory near Cologne was rejected (it cost more).

The committee puts some of the most expensive projects on ice. HERA, a high energy proton-electron collider for the DESY subnuclear physics laboratory at Hamburg which would cost DM 600 million (in two parts) is recommended "in principle", but a decision should await developments in superconducting technology; and the question of a relativistic heavy ion accelerator for GSI Darmstadt (DM 190 million) is to "remain open".

The latter raises an interesting question in physics: will relativistic heavy ion collisions lead to entirely new states of nuclear matter (as in neutron stars) as some nuclear physicists predict, or will it just lead to a mass of high energy but otherwise familiar fragments? The feeling in Germany is that some low cost test of this question should be made first, for example, by feeding heavy ions into the intersecting storage rings at CERN. But if such experiments were done, German thinking goes, it should be *outside* CERN's budget or there will be no money for LEP.

On this issue, on HERA, and on others such as the proposed spallation neutron source (where Jülich and Karlsruhe have competing proposals) and the deep drilling programme of the Deutsches Forschungsgemeinschaft, Germany will also be looking for increased international collaboration. The European synchrotron radiation source, however, proposed by the European Science Foundation, is said to be "not of high priority". Robert Walgate

Madrid conference

Detente denied

Western leaders have failed to react "realistically" to Mr Brezhnev's latest proposals for detente, a *Pravda* editorial complained last week. The proposals, put forward at the party congress last month, covered three main issues: the extension of "confidence building measures" (CBMs) to all Soviet territory west of the Urals, and the possibility of introducing similar measures in the Far East, special negotiations on the Persian Gulf and Afghanistan, and a moratorium on medium-range nuclear weapons in Europe.

The party congress coincided with the closing weeks of the Madrid review conference on implementation of the Helsinki Final Act. CBMs, as specified in the Final Act, cover notification of major military manoeuvres (more than 25,000 troops), exchange of observers for manoeuvres and also notification of major troop movements, though this last clause is left to the discretion of the participating states. Notification of manoeuvres, moreover, in the case of a participating state whose territory extends beyond Europe, is mandatory only for an area within 250 km of its European frontier. The suggestions that the Helsinki clauses on CBMs should extend as far as the Urals, came, initially, from the French delegates to Madrid, and were not greeted with enthusiasm by the Soviet delegation, until, on 23 February, Mr Brezhnev made them his own. Not surprisingly, Mr Brezhnev suggested some kind of reciprocity from the West. Since the whole of Europe falls within the CBM zone, this would presumably have to be of a qualitative nature, such as the discretionary notification of smaller-scale troop movements, or, judging from hints in the Soviet media, curtailing "provocative" broadcasts to the Soviet bloc.

Afghanistan, too, was an issue which the Soviet delegation at Madrid at first tried to block as "irrelevant" (because the Helsinki Final Act referred to security and cooperation in Europe). Mr Brezhnev's call for a moratorium on nuclear arms seems to reiterate the Polish proposal at Madrid for a European disarmament conference in autumn 1981. Unfortunately, none of Mr Brezhnev's peace proposals touched on the sensitive issue of declaration and verification of troop and armaments strength — the point on which all similar negotiations have broken down for the past 35 years. His suggestion of an "authoritative international committee" whose members "might include the most prominent scientists" of the countries concerned, which would "demonstrate the vital necessity of averting a nuclear catastrophe" has, however, been passed on to the international Pugwash council as a matter of urgency.

On the other aspects of detente, the

Moscow seminar threat

The Moscow Sunday seminars for Jewish refusenik scientists, which have been held (official harassment permitting) since 1973, now face a new threat. Last Friday, Dr Irina Brailovskaya was informed by the deputy chairman of the local municipal council that if she continued to host the seminars, she and her whole family would be liable to internal exile.

This official told Dr Brailovskaya that there had been complaints from the neighbours about the holding of the seminars. "We do not object to scientific meetings", he explained, "but the devil only knows what is going on here!"

The seminars have only recently resumed after a KGB clamp-down last autumn. Following the arrest of Dr Viktor Brailovskii, on 13 November, for several weeks the apartment was surrounded on Sundays by security police who turned away intending participants. Early in the new year, the seminars resumed on Saturday evenings, and then, a few weeks ago, reverted to the traditional Sundays.

In a recent message to Western scientists, Dr Brailovskaya said that her husband has been held in custody beyond the legal limit (three months) within which he should have been either charged or released. He is, she said, in a poor state of health, suffering from chronic hepatitis and a blockage of the gall duct.

auspices seem good for renewed East-West cooperation. President Reagan's assistant secretary designate on human rights, Dr Ernest Lefever, is reported to favour keeping such matters distinct from international politics. The several thousand Western scientists who signed personal moratoria pledges bound themselves to limit their contacts with the Soviet Union only until the end of the Madrid conference. However, at the meeting of the United Nations human rights commission in Geneva last week, a motion from the Canadian delegation was dropped from the agenda.

Vera Rich

Science education

Tap turned off

Washington

The Reagan Administration is proposing to eliminate all support at US universities for undergraduate and graduate science students at present provided by the National Science Foundation.

The cuts are part of a plan to phase out the foundation's responsibilities for science and engineering education, justified by the claim that these are of lower priority than the need to support basic research. But the proposed cuts may have been partially instigated by conservatives upset by some of the foundation's educational activities.

The cuts would eliminate several programmes introduced over the past few years, primarily at the direction of Congress. These include support for women, minorities and "talented individuals" in science, strengthening science teaching in middle and high schools, and efforts to improve public understanding of science.

The new Administration argues that such efforts are of low priority, and that they are too broadly spread and narrowly focused to have any significant effect. Science teaching has, however, always been treated better by Congress than by either the National Science Foundation itself or the Office of Management and Budget; and there remains a good chance that some of the money will be put back during the congressional review process.

In response to the report last year by the foundation and the Department of Education which drew attention to the growing "scientific illiteracy" among the US public and to concern over the lack of adequate technical graduates in fields related to energy and military technology, President Carter had proposed in January that the 1982 budget for the foundation should include an allocation of \$119.9 million for science and engineering education, a significant increase over the \$80.7 million allocated in the current year.

In contrast, the new Administration suggests that the current budget be reduced to \$64.7 million, and that support for science and engineering education next year be reduced to \$9.9 million. All of this sum would be used to meet existing commitments to students receiving fellowships.

Other cutbacks include the elimination of three major programmes: science education resources improvement (\$21.8 million in 1980), aimed at improving undergraduate instruction in two- and four-year colleges by upgrading instructional equipment and teacher competency; science education development and research (\$13.8 million) for developing new and more effective materials and methods of instruction in science and engineering; and science education communication (\$8.5 million).

Reaction to the proposed cuts has been sharp. Dr Alan Bromley, professor of physics at Yale University and president of the American Association for the Advancement of Science, warned that "to expect scientific and technological progress while abandoning efforts at improving science and engineering teaching in our schools is illogical and a disservice to the nation's interest". The association had previously announced that science education was to become a principal target of its own efforts over the coming year.

Dr Gerald Lieberman, dean of research at Stanford University, last week voiced the universities' anxieties: "In the past we have had to put little of our general funds to the support of graduate students in

Argentinians face charges

Dr Jose Federico Westerkamp, formerly professor of physics at the University of Buenos Aires, was arrested on 27 February 1981 during a police raid on the human rights organization CELS (Centro de Estudios Legales y Sociales, Centre for Social and Legal Studies). He was held in custody for a week, then released. Police investigations into his human rights activities are, however, continuing.

Dr Westerkamp, who had been active on behalf of human rights in Argentina for several years, was a co-founder of CELS, together with two lawyers — Dr Emilio Fermin Mignione and Dr Augusto Conte Mac Donell. The group, which was set up early in 1979 and legally registered in March 1980, investigates human rights abuses, including the defence of political prisoners and the gathering of data on "disappeared" persons. Dr Westerkamp, whose son Gustavo has been a political prisoner since 1975, has been particularly active on behalf of imprisoned and "disappeared" scientists and academics, and has on numerous occasions spoken on their behalf at international scientific meetings.

According to the Argentinian newspaper, *La Razon*, CELS members face charges under article 224 for possessing plans and diagrams of military installations, carrying a possible penalty of two to eight years imprisonment.

science and engineering, but it is likely that we will now have to reevaluate our whole policy in this area". Dr John C. Crowley, of the Association of American Universities, claims that the cuts would wipe out "one of the two functions for which the foundation was established" and "calls into question the basic purpose of the National Science Foundation".

At the foundation itself, where it has been frequently claimed that science education has received less support than basic research, officials admitted that the cuts were "hard to square" with the report presented to President Carter last autumn.

Federally-sponsored science education programmes have some powerful opponents on the right. Some argue that, in principle, the federal government should play a minimal role in curriculum development, leaving it to state and local schools boards; others have criticized the foundation's education activities in the past, for example in developing biology and social science courses which, they claim, reinforce "relativist" values by portraying man and woman as part of nature, rather than unique creations. The cuts in the science education budget seem to have been orchestrated by groups such as the conservative Heritage Foundation and are likely to receive support from conservative members of Congress.

David Dickson

Swiss nuclear power

Sanctions bite

Zurich

There is growing concern in Switzerland about uranium supplies for the nuclear energy programme. Indigenous deposits are too small to reduce foreign dependence, and the efforts of the chief suppliers, Canada and the United States, to prevent nuclear proliferation have led to acts which the Swiss consider to verge on blackmail. Although the pro-nuclear lobby maintains that the long-term prospects are "not unfavourable" and that the present supply hitch is merely a temporary inconvenience, there are signs that the country's independence, self-determination and neutrality will be threatened if nuclear power becomes a major element in its energy mix — just as today they are threatened by an overwhelming dependence on imported oil.

The nuclear plan — yet to go through parliament — foresees a capacity of about 7,000 MWe in conventional nuclear reactors by the beginning of the next century. With an average reactor life of 30 years, total requirements of uranium will be around 40,000 tonnes. A laborious search for indigenous deposits over the past 20 years has turned up several occurrences of interest, none of them economically viable.

The highest grade mineralizations occur in the Vallorcine granite in the Vallée du Trient (Valais) with average uranium contents of 0.1 to 0.2 per cent. But extensive exploratory drilling and trial shafts revealed reserves of less than 2,000 tonnes. Larger reserves are suspected in the upper Rhine valley near Trun (Graubünden), but there the average grade is too small (0.02–0.1 per cent). An unpublished government report in 1977, after an initial exploration costing 6 million francs (£1 = SF 4.33), concluded that the deposits in the Alps can be regarded only as emergency reserves. Normally, Swiss reactors will be completely dependent on imported fuel. It is this dependence which is beginning to cause some alarm.

The present dispute with Canada illustrates the issues. Since 1958, Canada has been the main supplier of natural uranium under a bilateral treaty. In 1977, however, the Canadian government refused permission for the transfer of yellow cake to the fuel fabrication plants in the United States which the Swiss utilities had already paid for. Since then, almost 1,000 tonnes of uranium have been blocked, and the Swiss utilities have been forced to buy alternative supplies on the international market — at far higher prices.

The reason for the blockade is the reluctance of the Swiss government to give in to Canada's demand for inspection and control rights with respect to some Swiss exports. Since the explosion of an Indian nuclear device in 1974, Canada has insisted

that every effort must be made to prevent sensitive materials, goods and technologies of "essentially" Canadian origin from being passed on to third countries without Canadian permission.

The Swiss government is in an extremely awkward position. The Bundesrat is being pressured from all sides and is still vacillating. The utilities consider Canada as their main long-term supplier and want the Canadian demands to be met at once, with as little fuss as possible. In fact, they have already negotiated the required amendments to the treaty, with rather furtive government backing. But this document has been gathering dust since 1979 because of its political explosiveness. The opposition by industrialists and politicians to a treaty modified under duress is expected to be loud and strong. The former see it as a completely unwarranted interference in the free market and object to controls not governed by Swiss laws. The latter see the loss of sovereignty and capitulation to outside pressure as matters of principle. Jurists have even suggested that Canada's failure to supply is a breach of the treaty which should be taken to the international courts — an interpretation which the Bundesrat has vehemently denied. And anti-nuclear members insist on asking awkward questions in parliament.

The general public, which is already rather sceptical of the future of nuclear energy, is confused. Industry and government experts maintain that the situation is under control and that it is only an initial

difficulty which, once mastered, will be followed by long-term guarantees of fuel supply. International cooperation and the direct financial participation of the utilities in uranium exploration will, they say, ensure that future requirements will be met. But the signs are hardly encouraging.

Two tendencies suggest that uranium mining and export will become increasingly controversial issues in the internal politics of the producing countries — the inevitable exploitation of lower grade ores, resulting in greatly increased environmental impact, and the increasing concern about the spread of nuclear weapons. The rapid development of nuclear power elsewhere will keep the world uranium market under pressure — and minor users like Switzerland at a constant disadvantage. If Switzerland's uranium supply today is being adversely affected by the explosion of a bomb nine years ago on another continent, what next?

Geoff Milnes

Indian community health

Cure for all ills

Lucknow

The Indian government's target of "health for all" by the year 2000 will not be met unless present policies are radically changed, according to a study group set up by the Indian Council of Medical Research (ICMR) and the Indian Council of Social Science Research (ICSSR).

Their report stresses that health should be recognized as a function of the overall development of society. An "alternative health care model" is presented which includes programmes for national economic growth with the objective of doubling national income per capita by the year 2000, full-scale employment at reasonable wages, improvement in the status of women, adult education, elementary education for all children between 6 and 14, welfare of backward classes, rural electrification, housing and organization of the poor. The report also suggests setting up a National Population Commission with the objective of stabilizing the total population at 1,200 million by the year 2050.

The "alternative health care model" encompasses promotive, preventive and curative aspects and combines "the valuable elements in Indian culture and tradition with the best elements of the Western system".

The report suggests that a Medical and Health Commission should be set up to look after the quantitative and qualitative aspects of health education. It recommends that drug production should be correlated with disease patterns with a bias towards the economic production of basic and essential drugs. On the research side, the important areas are delineated as parasitic diseases, problems of environment, sanitation and indigenous medicine.

Zaka Imam

More interferon cloned

The interferon genes of α and β interferon, produced respectively by lymphocytes and fibroblasts, have been cloned successfully in *Escherichia coli* at the University of Warwick, according to an announcement this week from the university. The project, led by Professor D.C. Burke, has been under way for the past three years. The Wellcome Foundation, one of several sources of finance, is also collaborating technically.

Although the two interferon genes have been successfully cloned in a bacterial plasmid, the genes have not yet been incorporated at plasmid sites where they are expressed, so as to produce the protein interferons.

Although at least three other groups (Biogen, Genentech and G.D. Searle) have successfully cloned interferons, doubts about the patentability of the techniques imply that being fourth is not necessarily to be robbed of reward.

For cloners of interferon, however, the most obvious prize for the immediate future is the cloning of gamma-interferon, otherwise known as immune interferon, which may have more interesting physiological properties than the two now in clinical trials.

Oil tanker hazards

Inaction at sea

Brussels

The recently published Liberian report on the *Amoco Cadiz* disaster draws attention to the belief widely held in shipping circles that 80 per cent of accidents at sea are due to human error. In March 1978 the tanker went aground off the coast of Brittany spilling 230,000 tonnes of oil. And now the captain is to get his licence back despite his failure to send out a distress signal at the earliest opportunity.



Amoco Cadiz — what happens next time?

The incident prompted the European Commission to spend two and a half years in preparing a set of proposals on how such accidents could be prevented and cleaned up in the future. The proposals were published last June, and despite pressures from the French government they run the risk of never being accepted.

The most controversial issue concerns a requirement that member states should identify sub-standard ships using their own ports and order deficiencies to be rectified. Now that Greece has joined the European Community, 30 per cent of the world's shipping falls under Community jurisdiction. A recent debate in the European Parliament on the idea that the port state should check the standards of shipping ended in uproar, with the Greek representatives defending themselves against accusations that most defective ships were Greek-owned.

Conventions on shipping safety have already been established by the Inter-governmental Maritime Consultative Organization (IMCO), but signatory countries have been slow to enforce them. The Commission's directive would make the IMCO conventions mandatory.

Dissatisfaction at the slow progress of both IMCO and European Commission proposals led France to hold a conference on shipping safety and pollution in

December last year. The Paris conference, which was attended by Spain, Portugal and Nordic countries as well as the European Community, set up a committee, which has been meeting in the Hague, with the declared aim of drawing up rules on the obligations of port states which would be acceptable to all European countries.

While these deliberations go on, the European Commission has made still more proposals, intended to stand more chance of being adopted quickly. One directive, already adopted, calls for all relevant data on tankers and their movements to be

stored on a computer. Another requires details of the contingency plans of all member states to be computerized, and the Commission wants to run training schemes and equipment testing programmes. Many such precautions are already in operation in the United States, causing operators with sub-standard ships to concentrate on European trade.

Whatever the various European Community bodies decide, nothing much is likely to happen for some time. The United Kingdom and Denmark still refuse to countenance giving the Community competence in this field.

In the meantime, France is still pressing for the £670 million estimated to be the cost of damages caused by the *Amoco Cadiz*.

Jasper Becker

US animal research

Stricter safeguards

Washington

Although the United States has so far avoided violent demonstrations against the use of animals in research such as those in Canada and the United Kingdom, public pressure for stricter safeguards is mounting. Three separate bills have been introduced in the new session of Congress. One, first introduced by Representative

Frederick Richmond in 1979, would permit the withdrawal of public funds from any research project in which laboratory animals were judged to be used unnecessarily.

The research community, and officials at the National Institutes of Health (NIH), are dismayed at the prospect of new regulations. At a meeting last month convened by NIH, scientists and animal welfare groups discussed the feasibility of the wider use of *in vitro* testing methods and mathematical modelling as alternatives to the use of live animals. Dr William Raub, associate director at NIH, said that he would approach the Office of Science and Technology Policy (if it survives under the new Administration) to suggest a regular forum for the discussion of alternatives to animals.

This proposal should at least satisfy Congressman George Brown, previously chairman of the House of Representatives Science and Research Committee, who came out in favour of more open discussion between scientists and their critics after a flood of correspondence supporting the Richmond bill. One of Mr Brown's objectives is to avoid trouble such as that in Canada, where radical anti-vivisectionists last year broke into National Research Council laboratories, setting free laboratory animals.

In the United States, protests against the widespread use of laboratory animals have taken the form of lobbying, led by groups such as United Action for Animals, based in New York.

One result so far stems from the campaign last year by more than 400 animal welfare groups against the Draize irritancy test for cosmetics, in which liquids are dropped on rabbits' eyes to test for inflammation. Revlon, one of the largest cosmetic manufacturers, was singled out for criticism and last November gave a \$750,000 grant to Rockefeller University to investigate cell culture alternatives.

The passage of new legislation will not, however, be straightforward. Present legislation goes back to the 1960s and is concerned with the care of laboratory animals, not with the tests to which they are subjected. The Scientists' Center for Animal Welfare is pressing for the wider adoption of a procedure already introduced by the Veterans Administration in which research proposals are vetted in advance to ensure that appropriate procedures are followed in the use of laboratory animals.

Mr Richmond's bill, the front runner in Congress, would establish a centre for the development of methods of research and testing not involving the use of live animals. Supporters of the bill say that such a centre could operate in much the same way as the National Toxicological Program, introduced three years ago, which coordinates research on toxic substances in the major federal agencies. The centre would have power to direct between

30 and 50 per cent of NIH funds for animal-based research into the development of alternative testing methods.

Critics, such as the National Society for Medical Research, argue that the bill would give federal regulators "potentially catastrophic" powers and that scientists use alternative methods voluntarily when they are available. Feelings on both sides run high, however. NIH officials, previously cool to the demands of the animal welfare groups, seem now to be hoping that if they can bend with the wind, legislation will be avoided.

David Dickson

UK animal research

Legislation unlikely

British legislation on laboratory animals is hanging fire until the draft of a Council of Europe convention is agreed, supposedly in May. The government has said that new legislation will not be based on the amended version of the private member's bill first introduced into the House of Lords by Lord Halsbury last session and reintroduced this session. Members of the Lords committee that sat to amend that Halsbury bill now fear that further delays, due to tardiness in the Council of Europe or in drafting a new bill, could jeopardize agreement reached by opposing sides during the committee's deliberations.

The government hopes to avoid passing legislation clashing with the convention's aims. British legislation, however, is bound to be more stringent than the lowest common denominator between the Council's member states laid down in the convention. The latest available draft contains nothing that contradicts the Halsbury bill or that is likely to give the government's bill drafters pause.

Whether further amendments will cause problems remains to be seen. But the government's objections to the Halsbury bill are clearly on other grounds. During the second reading debate before Christmas, Lord Belstead for the government said that the bill leaves too much to the discretion of the Home Secretary and invests too many powers in an advisory committee that would be set up to review regularly ethical matters and the replacement and use of laboratory animals.

Lord Halsbury is far from content. He suspects that the main objections to his bill have come from Home Office inspectors who fear that an advisory committee, along the lines laid down in the bill, would disrupt their well-tried methods of working. Despite government assurances that the original timetable still stands, Halsbury fears that delays in the Council of Europe could push it back. He plans to keep the issue alive by pushing his amended bill through the committee and report stages in the Lords in May even though it is now clear that it will never get a reading in the Commons.

Judy Redfearn

Commonwealth meeting

Food for reserves

Dacca

Moves to bolster the dangerously low level of world food stocks were the main outcome of a meeting of Commonwealth ministers for agricultural and rural development held in Dacca, Bangladesh, on 11-13 February. The ministers agreed that by mid-1981 negotiations should be completed to create a new International Grains Arrangement aimed at providing some of the 500,000 tons of cereals needed annually to replenish the International Emergency Food Reserve.

This was the first full-scale Commonwealth ministerial meeting of its kind, and was attended by 21 of the 44 member countries. The Commonwealth is made up of 20 nations showing a food deficit, and 21 with a surplus in food production. The remaining 3 nations, including host-nation Bangladesh, are on the verge of becoming self-sufficient.

The ministers urged quick action from international funding agencies, especially the World Bank and the International Fund for Agricultural Development, which are concentrating on aid to the poorest countries. Extra resources were called for to enable the Commonwealth Fund for Technical Cooperation to give more aid to developing Commonwealth countries.

M. Kabir

Romanian agriculture

Research expansion

President Nicolae Ceausescu of Romania called last month for greater research effort in agricultural sciences. He was addressing the 11,000 delegates at the Second Congress of Workers in Agriculture, including experts on agriculture as well as workers from cooperatives, collectives and the small residual private agricultural sector.

Although, according to Angelo Miculescu, the Minister of Agriculture, agricultural production went up by 26.4 per cent during the 1976-80 Five Year

Plan, President Ceausescu had previously pointed out, at a working meeting on agriculture in January, that targets during the past five years had not been fulfilled, and that "resolute" measures would still be needed. Such measures, he has now told the congress, must include expansion of research on plant and animal genetics, soil quality and, for the economists, the logistics of a switch from the traditional farming structure of the countryside, to the development of agro-industrial complexes. The latter scheme, by which food-processing plants are located in the countryside, not only eliminates the need of transporting perishable foodstuffs to the cities for processing, but provides additional jobs to reduce the drift to the cities.

The new research priorities will be met from the present budget, apparently by switching funds from other sectors.

In all these changes, said President Ceausescu, the Academy of Agricultural Sciences in Bucharest must have an increased involvement both in planning and coordination. The Academy, and, indeed, the whole educational system, he continued, must train more agricultural specialists.

Higher education in the agricultural sciences, however, is not to be open to all. President Ceausescu says that only workers or former workers in agriculture may go to university to study agriculture. The practice by which a young city-dweller can go to an agricultural institute simply to qualify for a clerical job nominally connected with agriculture is to stop.

Similarly, agronomy centres where specialists are trained are to be reorganized. At present, future experts are trained at "picturesque" centres, whose sites are frequently chosen for reasons other than agricultural suitability. At one such centre near Bucharest, for example, experts are trained in handling tractors in hilly areas, although the centre itself is remarkably flat. The reason is simple, said the president. "Comrades from the Ministry of Agriculture" want the centre there, because it is a convenient place for them to go and make speeches.

Vera Rich

Romanian workers — due for some changes



CORRESPONDENCE

Genesis of genetics

SIR — I find the current interest among biologists in the findings of Górczynski and Steele^{1,2} somewhat surprising, since a report of Lamarckian inheritance was published about 3,000 years ago. It is written that:

"And Jacob took him rods of green poplar, and of the hazel and chestnut tree; and pilled white strakes in them, and made the white appear which was in the rods. And he set the rods which he had pilled before the flocks in the gutters in the watering troughs when the flocks came to drink, that they should conceive when they came to drink. And the flocks conceived before the rods, and brought forth cattle ringstraked, speckled, and spotted."³

Were these findings ever reproduced? If so, would it not be more appropriate to speak of Jacobinic genetics?

KIRSTEN FISCHER LINDAHL

Basel Institute for Immunology,
Basel, Switzerland

1. Górczynski, R.M. & Steele, E.J. *Proc. natn. Acad. Sci. U.S.A.* 77, 2871-2875 (1980).
2. Górczynski, R.M. & Steele, E.J. *Nature* 289, 678-681 (1981).
3. *Genesis* Ch. 30, v. 37-39.

Ethics of genes

SIR — It was with considerable surprise and no little confusion that I read Richard Dawkins' letter on genetic determination (*Nature* 12 February, p.528). In his commendable desire to dissociate himself from the National Front, he has left me totally perplexed about his actual views of the relation between genotype and phenotype. Near the end of his letter, he associates himself with the views of S.J. Gould, that the genetic basis of IQ is "trivially true, uninteresting, and unimportant". Yet earlier in the same letter, he says that genetics is sort of relevant since we may need to "fight all the harder" against genetic tendencies. But in his book *The Selfish Gene*, Dr Dawkins wrote that we are "robot vehicles blindly programmed to preserve the selfish molecules known as genes" (preface) and that these genes "swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world . . . manipulating it by remote control. They are in you and me; they control us body and mind" (p.21).

It really is very vexing. Just as I had learned to accept myself as a genetic robot and, indeed, felt relieved that I was not responsible for my moral imperfections, Dr Dawkins tells me that, after all, I must try hard to be good and that I am not as manipulated as I thought. This is a problem I keep having in my attempt to understand human nature. Professor Wilson, in his book on sociobiology, assured me that neurobiology was going to provide me with "a genetically accurate and hence completely fair code of ethics" (p.575). I was

euphoric at the prospect that my moral dilemmas at last had a real prospect of resolution, when suddenly my hopes were dashed by an article in which Professor Wilson warned me against the naturalistic fallacy (*New York Times* 12 October, 1975). You can imagine my perplexity. I do wish I knew what to believe.

Perhaps I am just asking for that foolish consistency which Emerson tells us is the hobgoblin of small minds. But I see that Dr Dawkins himself is uncertain. I can only echo the question he asks in his letter. "Where on earth did the myth of the inevitability of genetic effects come from? Is it just a layman's fallacy, or are there influential professional biologists putting it about?"

ISADORE NABI

Museum of Comparative Zoology,
Harvard University, Cambridge,
Massachusetts, USA

Cancer causation

SIR — The Epstein-Swartz article which appeared in *Nature*¹ (and was essentially identical to the manuscript these authors presented at the autumn meetings of the American Public Health Association in Detroit) will elicit a variety of critical responses from those in the scientific community who are familiar with the specific aetiological links between cancer and the environment.

But here I would like to point out one major fallacy in the Epstein and Swartz piece — their premise (that lifestyle factors have been overstated in discussions of the aetiology of cancer) is based largely on the assumption that occupationally-induced cancers are more prevalent than previously thought. Epstein and Swartz criticize Dr Peto for dismissing "recent estimates of the importance of occupational carcinogens in a report by the US Public Health Service as exaggerated, unsound and unreasonable". This alleged "report" concluded that "as much as 20% or more" of cancers in the near term and future may reflect past exposure to specific carcinogens in the workplace.

Epstein and Swartz attempt to build up the credibility of this estimate by stating that the report was "prepared by nine named and internationally recognized experts in cancer epidemiology, statistics and carcinogens from three federal research agencies" and go on to add that "there is no basis whatsoever for recent unsubstantiated allegations by Peto and others that all or most of the authors of the government report have disowned or rejected it or its conclusions".

Epstein and Swartz are in error on two points in their discussion of occupational cancer and the so-called "estimates paper". First, nobody except those two authors and perhaps some regulators who wished to bolster their requests for more legal control over

occupational chemicals has ever cited the "20%" estimate in a serious, scientific context. Second, according to a detailed article by Ruth B. Schwartz in our publication *ACSH News and Views*² "7 of the 10 scientists responsible for these statistics believe they may be incorrect. Only one scientist defends them as reasonable". Two could not be reached for their opinions. Ms Schwartz's conclusions were based on her in-depth discussions with the various "contributors".

Perhaps the most succinct evaluation of the "estimates paper" on which Epstein and Swartz rely so heavily was made by Sir Richard Doll: "I regard it (the 'estimates paper') as scientific nonsense".

ELIZABETH M. WHELAN

American Council on Science and Health,
New York, USA

1. Epstein, S.S. & Swartz, J.B. *Nature* 289, 127-130 (1981).
2. Schwartz, R.B. *ACSH News and Views* Nov/Dec (1980).

Badger debate

SIR — Dr Flowerdew's recent letter (*Nature* 26 February, p.742) makes me wonder what purpose the Mammal Society intended to serve by engaging in what can indeed now be called "the badger controversy". Dr Flowerdew says that the society spoke out because my report provided a "one-sided" interpretation of the evidence ("biased" in Dr Harris's letter of 11 December; *Nature*, p.532), and because my "categorical" conclusions and recommendations did not take sufficient account of "the complexity of the problem". But in Dr Harris's letter it was also said that the society was intervening because it was necessary to correct certain "factually misleading statements" that I had made. Yet neither letter adduces any new "fact". Those that they use are the ones which I had assembled. Nor did either correspondent spell out any "anomalies" in the story other than those which I had considered. Dr Flowerdew now says that there is no unequivocal evidence that the gassing policy will produce a long-term solution, by which I presume he means stamp out the disease. But no one has ever claimed that it would. Above all, the Mammal Society, on whose behalf Drs Flowerdew and Harris have written, does not propose any alternative policy which the Government could pursue in order to lower the level of the reservoir of the bovine tubercle bacillus in the affected areas of the South West — the existence of which, as Dr Plowright pointed out (*Nature* 1/8 January, p.8), they first denied, but which in Dr Flowerdew's present letter he, or the Mammal Society, now seems to accept.

I intended no "attack on the Mammal Society and some of its members" — Dr Flowerdew's words — when I asked on what scientific authority the society based its presumed criticisms. The authorities on whom

I myself relied were first, the many scientists and veterinary surgeons from the Ministry of Agriculture, Fisheries and Food whom I cross-examined during my enquiry; second, the three internationally-known students of animal disease whose help I acknowledged in my report — Sir William Henderson, FRCVS, FRS, Dr L. Goodwin, FRCP, FRS, and Dr W. Plowright, DVSc, FRCVS; and third, Professor T. McKeown, FRCP, well known for his studies of trends in tuberculosis and other diseases; the authorities in the Medical Research Council and Department of Health and Social Services who are concerned with tuberculosis; and Dr B.R. Cook, until recently with the New Zealand Ministry of Agriculture.

The officers of the Mammal Society may believe that their society is exempt from the usual convention that no scientific society has the licence to make corporate statements of a scientific nature on behalf of all its members. But if it is exempt, one is entitled to ask — indeed, one is in duty bound to ask, since we are talking about a matter of animal and public health that is of international concern — who were the authorities on whom the Mammal Society leant, and whose counsel was denied me. Dr Flowerdew bluntly declares that the officers of the society were responsible for the statements that have appeared over his and Dr Harris's name. But apart from Dr Frazer, not one of the officers listed by the society is shown as having either a medical or veterinary qualification, and while some of its members, in particular Dr Neal, are well-known badger naturalists, none has written — unless in some unquoted journal — either on TB in cattle or on disease in badgers. Indeed, apart from Dr Neal's writings, I can find only one paper on badgers by any officer of the society, and that is a note by J.F.D. Frazer on badgers in Kent. The society's publications appear in the *Mammal Review* and in the form of "notes" in the *Journal of Zoology*, published by the Zoological Society of London.

Four of some 120 papers that have appeared in the lifetime of the *Review* concern badgers, and only seven deal with pathological matters (mainly ectoparasites — none deals with TB or with diseases in badgers); while only 15 of 311 "notes" concern badgers, and not one TB. In the circumstances, I find it more than a little strange that Dr Flowerdew suggests, presumably on behalf of his society, that the Ministry of Agriculture, Fisheries and Food should establish a Scientific Advisory Group "containing a small number of independent scientists such as wild-life epidemiologists and statisticians" — he offers no names — to "analyse the data in detail" for the ministry's Consultative Panel on Badgers and Tuberculosis, with the unfortunate imputation that in the view of the Mammal Society the ministry's scientific and veterinary officers are incompetent, and that their analyses of the data they collect cannot be relied upon.

LORD ZUCKERMAN

University of East Anglia,
Norwich, UK

SIR — Why has the Mammal Society (*Nature* 26 February, p.742) not attempted a reasoned reply to the criticisms contained in the letters by Dr Plowright (*Nature* 1/8 January, p.8) and myself (*Nature* 22 January, p.218), and by Zuckerman in his further article (*Nature* 19 February, p.628)?

If it is now accepted that the use of gassing

by the ministry has substantially reduced herd breakdown in affected areas, its continuation is clearly indicated. There does not at present appear to be any other way of tackling this source of infection. Reduction of badger densities in affected areas not only reduces the chance of infecting cows because there are fewer badgers around, but offers the best hope of greatly reducing and possibly, as more is learnt, substantially eliminating tuberculosis from the badgers themselves. In this respect badgers are fortunate in that there is now no major external source of infection and the TB monitoring of dairy herds gives early warning of the presence of infected setts in a locality.

In their earlier letter the Mammal Society, while suggesting that "other more subtle factors may be significant", did not dispute that density might be of importance in the spread and maintenance of TB in the badgers. They have now gone back on this, stating that "there is no unequivocal evidence that TB in badgers is density-dependent". This statement is based on the maps of badger density and herd breakdowns included in the Zuckerman report (p.68,70) which, they say, show that "TB is not only prevalent in areas of high density, but also in areas such as South Dartmoor where the badger population is very low". Had they referred to the subsidiary reliability diagram relating to the badger density map they would have found that the cluster of infections in South Dartmoor falls in an area classified on the reliability diagram as "mainly guesswork". They have completely ignored Zuckerman's discussion in his article of the statistical laws governing the propagation of infectious diseases, and his deduction that the establishment and maintenance of TB among badgers in an area is likely to be strongly density dependent, from which it follows that, as a matter of long-term policy, badger population densities should not be allowed to become excessive, as they certainly appear to be in parts of the South West at present.

The reactions to our other criticisms are similarly irresponsible and misleading. They dismiss Dr Plowright's letter with the comment that it "presented no new information nor did he answer any of the points we raised". The main point of my letter was to cast doubt on the Mammal Society's suggestion that the apparent success of the gassing campaign might be due to the decline of TB throughout the country. They appear now to have partially, but not wholly, conceded this point. They make much of the large decline from 5.5 to 3.2 per cent of herd breakdowns in Cornwall from 1975 to 1976 "highlighted" by my graph, and claim that it could not have been wholly produced by the gassing campaign. That there was a marked rise all over the country in 1975, followed by a corresponding decline in 1976, was clearly shown by my graph, and was pointed out in my letter, but this does not preclude that a large part of the Cornish decline was due to the removal of infected badger groups, some of which may have been done by farmers without the cooperation of the ministry. In any case, if the Mammal Society wished to comment on the irregularities of the Cornish data they might better have directed attention to the disturbing rise of infection between January 1977 and September 1979 which, however, was matched by a substantial fall between October 1979 and March 1980.

There are two further glaring distortions of the evidence in their current letter. The first is

my "contention", for which I gave reasons, that the 15 per cent of Cornish herd breakdowns attributed to badgers in the report was clearly much too low. That this contention was indeed correct was substantiated by the further information that Lord Zuckerman gave in his article. The Mammal Society accept this, but by lumping together the figures for 1974–75 and 1976–78 they conceal the progressive reduction in the proportion of "unknowns" and were able to claim that "since 1974 badgers were still only believed responsible for a minority of breakdowns".

A similar concealment of the full evidence occurs in their comment on the origins of infection from specific badger setts. Zuckerman reported that he had been informed that for the whole of Cornwall 72 per cent of the 51 outbreaks definitely attributed to badgers were within 2 miles of infected setts and 33 per cent were within half a mile, and that for the West Penwith area the corresponding percentages were 97 and 47. All that they state is that for the whole of Cornwall 28 per cent (100 minus 72) were attributed to setts more than 2 miles away and that badgers rarely travel this distance in areas of high density. A glance at the herd breakdowns map (p.20) of the report shows that the West Penwith area, which covers the western tip of Cornwall, contains the major pocket of infection and was clearly the source of the majority of the herd breakdowns.

Finally, in connection with their long and almost unintelligible argument on the effects of the moratorium on gassing, may I make it clear that neither Zuckerman nor I ever argued that "TB in badgers and cattle immediately increased as a result of the moratorium"; we merely reported the latest available data. Obviously there are time lags. The full effects of the moratorium cannot be properly assessed until data for a further year are available. Also in this connection, what is one to make of the statement: "This . . . totally ignores the likelihood of natural cyclic trends in TB prevalence, as discussed in our original letter and the reply by Dr Yates"? There is no mention of cyclic trends either in their letter or in my reply!

FRANK YATES

Rothamsted Experimental Station,
Harpenden, UK

SIR — You have recently drawn attention to the controversy over the badger-gassing issue.

The Zuckerman report shows a rise and fall of tuberculosis in the South West and also in the rest of England from 1974 to 1980. Some factor, common to the South West and the rest of England, must therefore be sought. I believe that factor is one which the Zuckerman report carefully avoids — the resumption of health-certification of the Irish cattle imports in June 1976 after nearly two years of importing TB-infected cattle into the United Kingdom during the Irish veterinary inspectors' strike.

The reluctance of the Animal Health Authority to suspend Irish cattle imports at that time contrasts sharply with their present enthusiasm to destroy badgers on scant evidence.

PETER H. ROBERTS

Compassion in World Farming,
Petersfield, UK

NEWS AND VIEWS

Propagating rifts in seafloor spreading patterns

from Tanya Atwater

SPECIALISTS from diverse geological fields recently gathered together to discuss mutual studies in the Cascadia region of western North America. One of the highlights of the meeting was the showing of a computer simulation of propagating rifts in the Juan de Fuca region by Richard Hey of the Hawaii Institute of Geophysics. Ever since Vine and Wilson¹ applied their seafloor spreading description to the well mapped magnetic anomalies of this region, enthusiasts have been cutting and pasting coloured maps, trying to explain the oblique fractures and other complexities in the pattern. For several years Hey and his colleagues have been developing and refining the notion of propagating rifts^{2,3} to explain data collected along the relatively remote Galápagos spreading centre. When they applied the concept to the Juan de Fuca anomalies, the many complexities became easily understood and predictable.

Probably the most convincing evidence for seafloor spreading, continental drift and plate tectonics lies in the configuration of linear magnetic anomalies from the sea floor. Vine and Matthews⁴ and Vine⁵ showed these anomalies to be indicators of symmetrical additions of new material along mid-ocean ridges and Wilson⁶ and Vine and Wilson¹ showed that these anomalies are offset along transform faults and their fossil traces, fracture zones. The now classic pattern of anomalies and offsets is that shown in Fig. a. This pattern is constructed assuming that the transform

offset was inherited from the original shape of the spreading boundary and that it constitutes impenetrable end points for the two ridge segments which abut it.

The simple pattern of Fig. a is a useful model both for explaining the gross configurations of seafloor features and for determining world plate motions. For fine details, on the other hand, the pattern is an oversimplification. Anomaly profiles commonly show occasional local sections of extra or missing space compared with predicted symmetrical patterns. These perturbations have generally been attributed to small 'ridge jumps' at which time the centre of spreading suddenly shifted sideways to a new location, breaking a section off from one plate and adding it to the other. In interpretations of sparse magnetic profiles, the detected spreading centre jumps have generally been assumed to be instantaneous breaks across new ground. More dense surveys have shown this not to be the case, however. Rather, the new ridges appear to tear through the crust at rates comparable to spreading rates, 'propagating' forward from their tips and gradually overlapping and extinguishing the previously existing spreading centres.

The magnetic pattern created by such a propagating rift is most easily visualized in its stepwise form, as shown in Fig. b after Hey². At regular intervals, the lower right-hand rift propagates past the transform fault, taking over the spreading for a new length of ridge and establishing a new

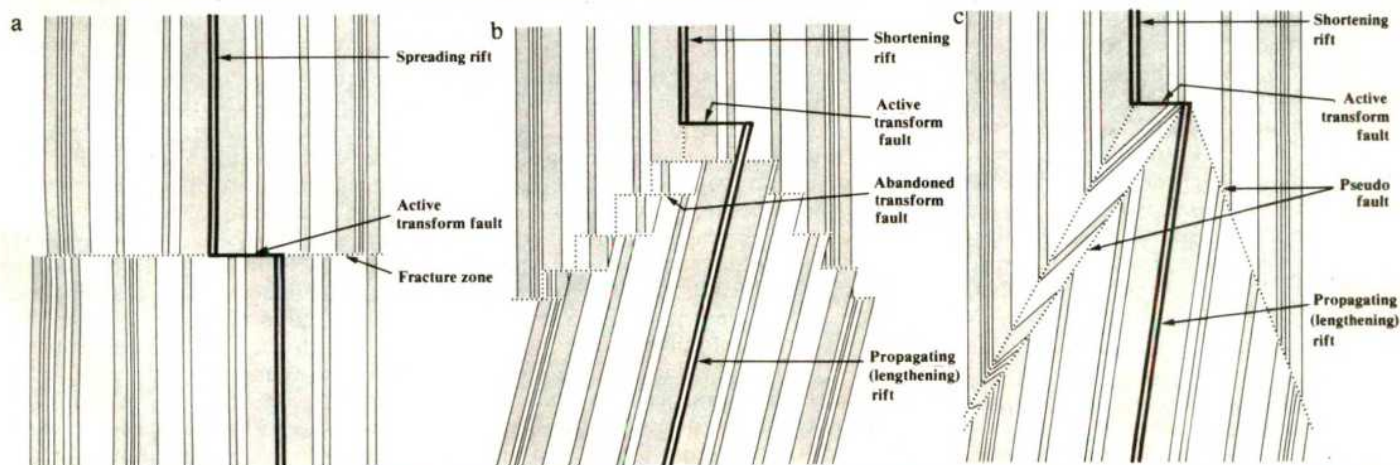
transform fault at its tip.

The third figure, after Hey *et al.*³, shows the anomaly pattern formed by a continuous, slow propagation of the lower rift and continuous shifting of the connecting transform zone. The oblique V-shaped traces left in the anomaly pattern appear at first sight to be fault lines along which the anomalies were broken and offset after their formation. Hey² notes that, in fact, no fractures ever occurred in these directions. They are inherited patterns which he calls 'pseudofaults'. V-shaped patterns like these are very clear in the Juan de Fuca magnetic surveys⁷, in the Surveyor surveys further offshore⁸ and in the Galápagos surveys^{3,9}. In the Atlantic, magnetic anomalies are much less clear because of slow spreading rates, but V-shaped patterns have long been noted in the topography near the Reykjanes ridge¹⁰ and at mid latitudes¹¹.

Another phenomenon clearly related to the propagators is the occurrence of ferro basalts and related high-amplitude magnetic anomalies¹². These unusual basalts erupt primarily near the growing rift tips and may help us to understand the special environment in which these rifts occur.

The reason why one rift propagates to supercede another is not clear. It is common for the new rift to have a slightly different trend than the old, suggesting this is the mechanism by which spreading centres adjust to changes in stress directions. However, the propagation of

Marine magnetic anomaly patterns generated by seafloor spreading at: (a) A stable spreading-transform fault system¹; (b) An oblique rift propagating northwards in discrete steps²; and (c) An oblique rift propagating steadily northwards³. Oblique 'fractures' mapped at the Juan de Fuca and other spreading centres may actually be pseudofaults as shown in (c).



rifts can be seen to accomplish other interesting results, as well. In the Galápagos area, for instance, the rift has 'jumped' south several times (via propagating rifts), thereby maintaining a location near the Galápagos hot spot. Another suggestive relationship is that the propagators commonly migrate outwards from the proximity of hot spots³.

The examples of propagators thus far documented occur at transform faults of small offset. However, Hey *et al.*³ suggest that all rifts begin as propagating cracks. The present East Pacific Rise in the South Pacific seems to have slowly grown, gradually replacing one that lies, now extinct, to the east¹³. Similarly, the original rifting of the Atlantic may have been progressive as suggested by the non-isochronous edge of the magnetic quiet zone¹⁴. A more recent version of continental rifting can be seen in the Gulf of Aden and Gulf of Tadjourah where the rift appears to have propagated westwards

Tanya Atwater is Professor of Marine Geophysics at the University of California, Santa Barbara.

into the African continent at a rate of about 30 mm per year¹⁵.

The propagating rift idea is really only a small modification to our understanding of seafloor spreading. Yet, it is an extremely satisfying one since it explains so many diverse 'imperfections' in the record in such a simple way, and it gives us a much better insight into the processes by which rifts form and readjust to changing conditions. □

1. Vine, F.J. & Wilson, J.T. *Science* **150**, 485 (1965).
2. Hey, R. *Earth Planet. Sci. Lett.* **37**, 321 (1977).
3. Hey, R., Duenebier, F.K. & Morgan, W.J. *J. geophys. Res.* **85**, 3647 (1980).
4. Vine, F.J. & Matthews, D.H. *Nature* **199**, 947 (1963).
5. Vine, F.J. *Science* **154**, 1405 (1966).
6. Wilson, J.T. *Science* **150**, 482 (1965).
7. See, Vine, F. for example, *History of the Earth's Crust* p.73 (1968).
8. Shih, J. & Molnar, P. *J. geophys. Res.* **80**, 4815 (1975).
9. Hey, R. & Vogt, P.R. *Tectonophysics* **37**, 41 (1977).
10. Vogt, P.R. *Earth Planet. Sci. Lett.* **13**, 153 (1971); Johnson, G.L. & Vogt, P.R. *Geol. Soc. Am. Bull.* **84**, 3443 (1973).
11. Rona, P.A. *Earth Planet. Sci. Lett.* **30**, 109 (1976).
12. Vogt, P.R. & de Boer, J. *Earth Planet. Sci. Lett.* **33**, 145 (1976).
13. Herron, E.M. *Geol. Soc. Am. Bull.* **83**, 1671 (1972).
14. Hayes, D.W. & Robinowitz, P.D. *EOS Trans. AGU* **56**, 450 (1975).
15. Courtillot, V., Galdeano, A. & LeMouél, J.L. *Earth Planet. Sci. Lett.* **47**, 144 (1980).

The detection of cellular transforming genes

from Peter W.J. Rigby

THE demonstration that the transforming genes of retroviruses are derived from cellular genes, designated proto-oncogenes¹⁻³, has considerable implications. If certain cellular genes are oncogenic when placed under the influence of viral control elements, perhaps other genes, or indeed the same ones, mediate chemical carcinogenesis by being expressed at high levels as the result of carcinogen-induced mutations. Recent work from the groups of Robert Weinberg (MIT) and Geoffrey Cooper (Harvard University) has shown that genes involved in chemical carcinogenesis can be detected by DNA-mediated gene transfer experiments.

Weinberg's group has been able to induce transformation by transfecting into non-transformed mouse fibroblast DNA from lines of mouse fibroblasts transformed *in vitro* by the chemical carcinogen 3-methylcholanthrene (MC)⁴. Efficient transfection occurs with only one such line, NIH3T3, the cells of which are presumed, for some reason, to be particularly efficient at taking up and expressing exogenous DNA. DNAs from fifteen transformed lines were tested and successful transfer of the transformed phenotype was obtained with DNA from five lines of MC-

transformed cells. High-molecular-weight DNA (~30 kilobases in length) was used and the efficiency observed was roughly 0.1 focus forming unit per µg of DNA (FFU µg⁻¹). Cell lines derived from the foci were subsequently shown to exhibit another parameter of transformation — the ability to grow in soft agar. Infection of the transformed cells with murine leukaemia virus failed to rescue a transforming virus, suggesting that transformation was not due to the activation of an endogenous retrovirus. The transformed phenotype could be passaged through a second round of transfection. Note that the DNA of the majority of cell lines tested could not transfer the transformed phenotype and, more important, neither could the DNA of some MC-transformed lines.

In a second report⁵, restriction endonuclease digestion of MC-transformed cell DNA was used to assess the number of transforming genes. A gene will be inactive in the transfection assay if an enzyme cuts within it whereas enzymes that cut outside the gene will have no effect. In each of four cases tested, *EcoRI* and *HindIII* abolished transformation whereas digestion with *BamI*, *XhoI* and *Sall* had no effect. It thus appears that the same gene is involved in transformation in each of the four cell lines. The transforming genes of other cell types, a rat neuroblastoma and a line of cells transformed by transfection with normal mouse DNA, exhibited

distinct patterns of restriction enzyme sensitivity. The lines showing the same pattern were all derived by *in vitro* transformation of mouse fibroblasts with MC. The DNA of many MC-transformed cells was not active in the transfection assay, so it seems likely that only a subset of MC-responsive sequences was being examined.

In this issue of *Nature* (p.261), Shih and his colleagues extend their observations by showing that the phenomenon is not confined to genes involved in the transformation of mouse fibroblasts. Cellular transforming genes can cross cell-type and species barriers — DNA from three out of four rat neuroblastomas induced by ethylnitrosourea transformed NIH3T3 cells as did DNA from a MC-induced mouse glioma. In each case the primary transfectants were tumorigenic and a second cycle of transfection could be demonstrated. Similar results were obtained with a variety of carcinomas, including the Lewis lung carcinoma of the mouse, a 7,12-dimethylbenzanthracene-induced mouse bladder epithelial carcinoma, a benzo-a-pyrene-induced rabbit bladder epithelial carcinoma and a human bladder epithelial carcinoma. The phenomenon was again not general, the DNAs of several neuroblastomas and carcinomas being inactive, and rescue experiments argued against the involvement of retroviruses. This paper provides direct evidence that mouse cells transformed by human carcinoma DNA have acquired human sequences. In DNA-mediated gene transfer experiments, cells generally seem to take up not only the gene that is being selected for (in this case the carcinoma transforming gene) but also any other DNA that happens to be in the calcium phosphate precipitate. Human DNA contains a repeated sequence family, known as the *Alu* family, that is not detectable in the mouse genome. Using a cloned *Alu* fragment as hybridization probe it was shown that the transformed mouse cells had also acquired many copies of the human *Alu* sequence.

Cooper's laboratory has reported similar results although there are important differences of detail⁶. Using DNA 30 kilobases or more in length (the same size as used by Weinberg's group), they obtained no transformation with DNA from MC-transformed cells of murine and avian origin or with DNA from corresponding normal cells. However, if the DNA was sheared to between 0.5 and 5 kilobases, transformation was obtained with both normal and transformed cell DNA. Cooper and his colleagues examined only one line of MC-transformed mouse cells so their failure to detect transformation with high-molecular-weight DNA does not conflict with the data from Weinberg's group, who found that only one-third of such lines scored positive in the transfection assay. The frequency at which Cooper's group observed transformation

Peter W.J. Rigby is a member of the Cancer Research Campaign Eukaryotic Molecular Genetics Research Group, Imperial College of Science and Technology, London.

with normal cell DNA, even after sonication, was extremely low and Shih and colleagues, observing such transformation, in fact at a higher frequency, regarded these data as of dubious significance. When Cooper *et al.* used DNA extracted from their primary transformants in a second cycle of transfection, they again observed transformation but at a much higher frequency (0.1–0.2 FFU μg^{-1}), comparable with that obtained by Shih *et al.* for the lines that they scored as positive. Cooper *et al.* have also performed virus rescue experiments which argue against a role for retroviruses and have used restriction enzyme digestion to show that there are at least two normal chicken cell genes capable of inducing transformation.

The following general picture emerges from these results. DNA from some chemically induced and spontaneous tumours will induce transformation even in high-molecular-weight form. In these cases the transforming gene is dominant and the genetic lesion causing transformation is likely to be *cis*-acting, for example an up mutation in the promoter controlling transcription of the gene. It is not possible to conclude that the gene detected is the primary target of the chemical carcinogen. In many other cases, however, the transformed phenotype is not transferable and the genetic lesion may therefore be *trans*-acting, for example a defect in a diffusible regulatory molecule. In these cases transformation may be induced if the DNA is sonicated to a small size before transfection. The model favoured by Cooper's group is that the sonication separates the transforming gene from regulatory sequences that normally control its expression. In the primary transfection, which is very inefficient, transformation occurs if the gene integrates close to a strong promoter. In a secondary transfection a high efficiency is obtained because the lesion is now *cis*-dominant.

This model depends on dosage. It says that genes which during normal expression are not deleterious, and may even be vital for normal growth and development, can be oncogenic when joined to regulatory elements which increase their expression. This model has received direct support from experiments by Vande Woude and his colleagues⁷. They have cloned a proto-oncogene, the murine cellular *src* gene, from normal mouse cells and shown that it does not transform when transfected into cells. If, however, the cellular *src* gene is joined to sequences from the 5' end of murine leukaemia virus which contain the strong viral promoter, the resulting recombinant DNA molecules transform efficiently.

Many questions remain. Clarification of the ability of normal cell DNA to transform and the effect of DNA size on transformation frequency is particularly important. While Weinberg's group have treated their data with great caution, they

are amenable to the interpretation favoured by Cooper and his co-workers, and one eagerly awaits the results of more detailed comparative studies. A second question is how many transforming genes are there and what is their relationship to the genes found in transforming retroviruses? The transfection assay should allow the detection of many transforming genes and Rapp and Todaro⁸ have published experiments suggesting that it will be possible to rescue genes capable of inducing a wide variety of tumours using retroviruses as vectors. The demonstration that the malignant phenotype of chemically transformed cells can be transferred by DNA transfection provides a ready assay which will allow the corresponding genes to be cloned. The availability in cloned form of genes

involved in chemical carcinogenesis will surely lead to enormous advances in our understanding of this process. Of critical importance is the demonstration that human transforming genes can be detected by transfection because the failure to isolate human transforming retroviruses has meant that we remain totally ignorant of the properties of the proto-oncogenes which are surely present in the human genome. □

1. Cold Spring Harb. Symp. quant. Biol. 44 (1980).
2. Hayman, J. *Gen. Virol.* 52, 1 (1981).
3. Bishop, J.M. *Cell* 23, 5 (1981).
4. Shih, C. *et al. Proc. natn. Acad. Sci. U.S.A.* 76, 5714 (1979).
5. Shilo, B.Z. & Weinberg, R.A. *Nature* 289, 607 (1981).
6. Cooper *et al. Nature* 284, 418 (1980).
7. Oskarsson *et al. Science* 207, 1222, (1980).
8. Rapp & Todaro *Proc. natn. Acad. Sci. U.S.A.* 77, 624 (1980).

Proline and protein procrastination

from Roger H. Pain

EVER SINCE it was realized that a polypeptide could not afford the time to explore all possible configurations before settling in the unique native conformation, the main thrust of work in the field of protein folding has been concerned with understanding how folding can take place so rapidly. The sequences of proteins are designed not only to ensure appropriate stability and biological function but also, and of equal importance, to direct the folding reaction along pathways which shortcut the majority of configurations and allow the native conformation to be reached in well under a minute. The broad outlines of how this happens are now fairly clear.

It is therefore strange, at first sight, to find that a mechanism currently the object of a substantial amount of effort is one whereby this highly sophisticated physico-chemical process is actually delayed by the molecular eccentricity of a few proline residues (see Freedman *Nature* 279, 756; 1979). The biphasic kinetics of refolding exhibited by many proteins have been explained by the general model



where U_S is an unfolded conformation which folds to the native state, N , with relatively slow kinetics because it has to undergo a prior rate-limiting change into the fast refolding form U_F . U_F contains prolines in the *cis* or *trans* isomers which

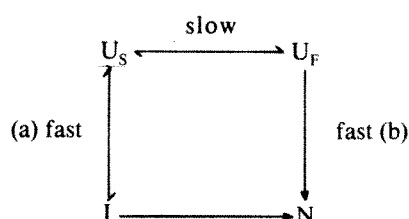
occur in the state N , but U_S is a form in which a proportion of non-native isomers have arisen by slow *cis-trans* isomerization. Thus, the kinetics of refolding are modified by a transformation which is itself almost spectroscopically invisible.

The evidence for this explanation is, in the words of Jullien and Baldwin (*J. molec. Biol.* 144; 1981), "strong but not yet complete". It has been supported by the key double-jump type of experiment. If state N is perturbed to state U_F by thermal, acid or solvent denaturation (jump 1), the kinetics of the refolding to N after removal of the perturbation (jump 2) will depend on the time allowed for the $U_S \leftrightarrow U_F$ equilibrium to become established. It is found that the slow phase becomes more pronounced the longer the time interval between the first and second jumps and that the activation energy is similar to that for proline isomerization. Furthermore, strong acid, which catalyses proline isomerization, also increases the rate of equilibration between U_F and U_S .

The above equilibrium is the simple model initially put forward in 1975 by Brandts, Halverson and Brennan (*Biochemistry* 14, 4953). However, it has always seemed intuitively possible that U_S could fold to a compact configuration, still containing the 'wrong' proline isomer, but then reorganize itself on subsequent isomerization to state N . Baldwin and his collaborators have shown that for ribonuclease this is true and, indeed, that the isomerization process can be speeded up in this folded state by as much as

Roger H. Pain is in the Department of Biochemistry at the University of Newcastle upon Tyne.

50-fold. The more general basic model for refolding is now



where I represents one, or a sequence, of intermediate states with improper prolines. Levitt has drawn attention to the consequence inherent in this model that the refolding route taken will be a function of the difference in conformational energy between the native state N and the similar though perturbed conformation containing the wrong proline isomer (*J. molec. Biol.* **144**; 1981). The higher the strain associated with the incorporation of incorrect proline isomers into the native conformation, the greater the relative importance of pathway (b) and the more the system will approximate to Brandts' original model. At the other extreme, if the difference in energy is low, pathway (a) will predominate and isomerization only be detectable insofar as the folded state I differs spectroscopically or in stability from state N. State I will be expected to be significantly populated under conditions favouring the 'native' state.

Levitt, on the basis of conformational enthalpy calculations, postulates three levels of strain energy resulting from the insertion of incorrect proline isomers, types I, II and III corresponding to energy differences of approximately 1, 10 and 30 kcal mol⁻¹. Types III and I correspond, respectively, to the two cases cited above, while in type II the folded state containing an incorrect proline would be relatively stable but significantly less so than the native state.

Jullien and Baldwin have taken up this classification and have suggested kinetic and environmental criteria whereby the three types may in principle be recognized experimentally. On the basis of spectroscopic and urea gel electrophoresis (Creighton *J. molec. Biol.* **137**, 61; 1980) investigations on about ten proteins, they find definite evidence for proteins with type II prolines only, although the existence of type I prolines cannot be ruled out. Thus, in any system some molecules refold fast, others slowly. Generalizations about globular proteins are notoriously prone to contradiction, however, and investigation of a wider range of proteins may reveal different behaviour.

Levitt's main contribution is to calculate the strain energies for each of the four prolines placed, as the incorrect isomer, in bovine pancreatic trypsin inhibitor (BPTI). The effect on the protein is shown to involve mainly local and relatively small shifts from the native conformation. He finds that the destabilization of the native

structure is 1 kcal mol⁻¹ for Pro 13 (therefore type I), 11 kcal mol⁻¹ for Pro 2 and Pro 9 (type II) and 33 kcal mol⁻¹ for Pro 8 (type III). On the basis of these assignments, one can calculate the proportion of BPTI molecules which will refold rapidly and Levitt obtains a value of 0.77. Jullien and Baldwin, who carry out a detailed experimental investigation of the refolding of a slightly modified BPTI, find that the relative amplitude of the fast folding $U_F \rightarrow N$ reaction is 75 per cent, in good agreement with the prediction. They also find one slow folding reaction which satisfies the criteria for a type II residue together with preliminary evidence strongly suggestive of one or two type I residues. Type III behaviour has not been found, although this may be due to greater conformational mobility in the modified BPTI which they study. The 'catalysis' of isomerization found experimentally has been studied by Levitt who calculates that states intermediate between *trans* and *cis*, with the peptide bond angle $\omega \sim 90^\circ$, have lower energies than the strain energies in

the *cis* form. The protein seems to be able to accelerate the rate of isomerization by applying a couple to the ω angle.

Taken together, these two different approaches provide more circumstantial evidence for proline isomerization as the explanation for the biphasic refolding of proteins. If the fast and slow phases do, in fact, correspond to the sequential $U_S \rightarrow I$ and $I \rightarrow N$ reactions where I is a compact configuration, and if most proteins turn out to fold as type II or type I, the proline quirk affects the exquisite efficiency of the folding process less than might have been initially feared. The polypeptide chain still collapses rapidly, as suggested by Creighton, to a compact form within which the prolines can sort themselves out at leisure. The collapse may be the most important feature in that it reduces the possibilities for entanglement between elements of different domains. Presumably the conformational requirements for proline residues in the folded protein outweigh the necessity for an overall rapid folding process. □

Defining a lectin

from Jan Kocourek and Václav Hořejší

Recent correspondence (see *Nature* **285**, 66; 1980) demonstrates that there is an urgent need to formulate a definition of lectins which clearly delineates the meaning of the term. However, Goldstein *et al.*'s proposed definition of a lectin as "a sugar-binding protein or glycoprotein of non-immune origin which agglutinates cells and/or precipitates glycoconjugates" seems to disregard some important points. In particular, because the role of lectins in living objects is unclear, it does not seem appropriate to define lectins by their *in vitro* biological activities. The agglutination of cells — a complex and poorly understood phenomenon — may, in the majority of cases, be a fortuitous effect of lectins. The definition of lectins should be based primarily on their most important physicochemical properties, that is, on their interaction with carbohydrates.

The inclusion into the definition of some *in vitro* activities of lectins, like agglutination of cells and precipitation of glycoconjugates, would exclude those lectins which have only one binding site and/or show some toxic or hormone-like activities. This is especially true in the case of the toxins of *Ricinus communis*

and *Abrus precatorius* which are structurally related to the 'agglutinins' and occur in the same sources so that their common evolutionary origin is obvious. Recent research shows that many of these agglutinins (including that of *R. communis*) may exert very similar properties to their toxic counterparts, if acting in cell-free systems (Barbieri *et al. Biochem. J.* **182**, 633; 1979).

As long as we consider lectins as recognition molecules and assume that they mediate information in various (but still largely unknown) physiological reactions, their definition should be based on only three distinct features which they all have in common. Lectins

- (1) have at least one binding site which can bind free or glycosidically linked carbohydrates: this interaction can be inhibited by low-molecular-weight sugars;
- (2) are not synthesized due to an immune response and do not have immunoglobulin structure;
- (3) do not show enzymatic activities towards carbohydrates to which they bind.

We thus suggest the following definition: lectins are sugar-binding proteins or glycoproteins of non-immune origin which are devoid of enzymatic activity towards sugars to which they bind and do not require free glycosidic hydroxyl groups on these sugars for their binding.

Jan Kocourek is in the Department of Biochemistry, Charles University, Prague and Václav Hořejší is in the Institute of Molecular Genetics, Czechoslovakia Academy of Sciences, Prague.

Chondritic meteorites: a changed and changing view

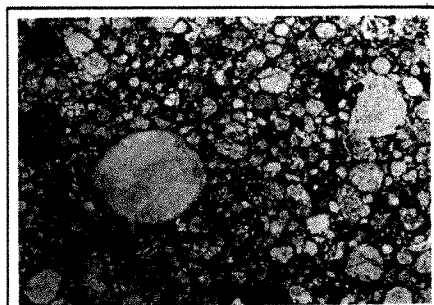
from R. T. Dodd

Two decades of intensive study of meteorites has drastically changed our view of their role in Solar System history. What R.A. Daly regarded in 1943 as fragments of a shattered tenth planet¹, we now recognize as samples of at least 20 (ref. 2) and perhaps more than 70 parent bodies³, most or all of which were no larger than Ceres, the largest modern asteroid (with a radius of 385 km)^{2,4}. Some of the meteorite parent bodies experienced partial melting and fractional crystallization to yield the diverse rocks that we know as iron, stony-iron and achondritic stony meteorites. Others escaped such differentiation and yield samples of rocks — chondritic meteorites — that formed during or very shortly after condensation of the solar nebula and underwent little physical and less chemical alteration thereafter. While the differentiated meteorites provide valuable insights into planetary evolution⁵, chondrites contain a unique record of the earliest stages in Solar System history, and some of them contain isotopically bizarre materials that may predate formation of the solar nebula and/or come from beyond it⁶. All meteorites are important, but the chondrites alone among accessible materials are survivors of the first steps in the 4.55 Gyr progression from interstellar gas to modern man.

Chondrites are more or less recrystallized agglomerate rocks (see the figure) whose bulk chemical compositions (ignoring H, He and other gases) resemble the composition of the Sun. Though all chondrites are, in this sense, chemically primitive, they show discontinuous variations of oxidation state and ratios of major non-volatile elements that divide them into three chemical classes and at least eight smaller chemical groups (see the table). There is general agreement that these variations reflect processing of chondritic material before or during its accumulation in at least eight parent bodies, though the nature of that processing remains unclear^{4,7,8}. There is also a consensus that the recrystallization evident in most ordinary and enstatite and some carbonaceous chondrites signifies the thermal metamorphism in^{4,9} or on^{8,10,11} the

surfaces of the parent bodies; the specific mechanism is again unclear⁴.

Much discussion of chondrites centres on the origin of chondrules, the millimetre-sized, drop-formed or irregular bits of rock from which these meteorites take their name (see the figure). Chondrules are present in most carbonaceous and all enstatite and ordinary chondrites, and their origin is highly controversial. Of the



Photomicrograph, in transmitted light, of the Manych ordinary chondrite. Chondrules, which comprise almost 90 per cent by volume of this meteorite, consist chiefly of olivine, pyroxenes and residual glass. The field of view is 16 mm wide.

many processes invoked to explain them — condensation of droplets from nebular gas, fusion of nebular dust by various means, shock melting of dust or rock on parent body surfaces, explosive volcanism and metamorphism — only the last two have no modern proponents, and only the last is clearly inconsistent with the properties of chondrules.

Debate over the origin of chondrules remains spirited, but new analytical methods (electron microprobe, neutron activation and isotope dilution analysis) have provided analyses of individual chondrules that place tight constraints on models. Chemical trends among chondrules in carbonaceous chondrites strongly suggest that these objects, like the refractory inclusions in such meteorites, formed in the condensing solar nebula, by condensation of droplets and/or fusion of solid condensates¹². On the other hand, chemical variations among chondrules in

ordinary chondrites suggest that they post-date condensation and formed through shock melting associated with the growth of parent bodies¹³. It thus appears necessary to appeal to different processes and/or times of formation⁴ to explain chondrules in carbonaceous and ordinary chondrites.

Chondrules in enstatite chondrites have received little attention to date, largely because of the perception that all such meteorites are metamorphic rocks¹⁴. Two recent papers on such chondrules^{15,16} are therefore welcome. They suggest that metamorphism is a less severe problem than believed, and that the chondrules in enstatite chondrites are remarkably complex objects that may require yet another explanation.

Leitch and Smith showed that chondrules in the Indarch chondrite contain two types of clinopyroxene (essentially MgSiO_3)¹⁵ and two types of olivine (essentially Mg_2SiO_4)¹⁶ that differ in minor element contents and cathodoluminescence colour. That these and similar phases in the Kota-Kota chondrite¹⁶ occur together but do not intergrade makes it unlikely that either meteorite experienced significant metamorphic alteration. In the standard classification of chondrites by composition and degree of recrystallization¹⁴, they probably belong to type E3 rather than E4.

Leitch and Smith (ref. 15 and see this issue of *Nature* p.228) observed that the distribution of the two pyroxenes and two olivines is very complex — some chondrules contain all four minerals — and hard to reconcile with a direct connection between chondrule formation and nebular condensation. They suggest that enstatite chondrites bring together materials from at least two nebular sources, whose history included condensation, formation of planetesimals, collisional destruction of these objects and accumulation of the resulting debris.

Many aspects of this scenario need further study, as its authors admit. It is important, for example, to characterize the fine material between crystals in chondrules, which Leitch and Smith assume to be dust and vapour deposits. This writer is also troubled by their explanation of droplet chondrules, which tend to contain either red- or blue-luminescing pyroxene (or orange or blue olivine) and are thought to sample different planetesimals. It is not clear why liquids generated by collisions should sample only one of the colliding objects.

R. T. Dodd is Professor of Mineralogy in the Department of Earth and Space Sciences, State University of New York at Stony Brook.

Abundances and chemical properties of chondrite classes

Class	Groups	Falls	Fe/Si	Atomic Fe ^o /Fe	Mg/Si
Carbonaceous	>4	33	1.6–1.8	0.0–0.1	0.90–0.92
Ordinary	3	314	1.2–1.6	<0.1–0.6	0.80–0.83
Enstatite	>1	11	1.1–2.0	0.7–0.8	0.63–0.68

Abundance data from ref. 8; approximate ranges of ratios from ref. 4.

The Leitch-Smith model for chondrule formation is at present skeletal, but their demonstration that essentially pristine enstatite chondrites exist is very important. It encourages those (for example ref. 16) who are attempting to study chemical variations among chondrules in these meteorites. It should also encourage other workers to move into this neglected corner of chondrite research.

Evidence that the chondrules in carbonaceous, ordinary and enstatite chondrites tell quite different stories illustrates very well the marked change in view alluded to at the beginning of this article. Daly¹ interpreted chondrites as samples of a planetary mantle and thus, by implication, chemically and physically differentiated materials. Had he proved to be correct, we might now know more than we do about the interior of our planet, but we would surely know much less about its origin. A shattered tenth planet is emotionally appealing, but it would tell us much less about the Solar System's history than 20 or

more smaller objects caught at various stages of internal evolution. That the meteorite parent bodies were numerous and varied might disappoint Daly, but it is very good news for planetary scientists. □

1. Daly, R.A. *Bull. geol. Soc. Am.* **54**, 401 (1943).
2. Wasson, J.T. & Wetherill, G.W. in *Asteroids* (ed. T. Gehrels) 926 (1979).
3. Scott, E.R.D. *Geochim. cosmochim. Acta* **42**, 1447 (1978).
4. Dodd, R.T. *Meteorites: A Petrologic-Chemical Synthesis* (1981).
5. Stolper, E., McSween, H.Y. Jr & Hays, J.F. *Geochim. cosmochim. Acta* **43**, 589 (1979).
6. Clayton, R.N. *A. Rev. nucl. particle Phys.* **28**, 501 (1978).
7. Larimer, J.W. & Anders, E. *Geochim. cosmochim. Acta* **34**, 367 (1970).
8. Wasson, J.T. *Meteorites, Classification and Properties* (1974).
9. Dodd, R.T. *Geochim. cosmochim. Acta* **33**, 161 (1969).
10. Heyse, J.V. *Earth planet Sci. Lett.* **40**, 365 (1978).
11. Hutchison, R., Bevan, A.W.R., Agrell, S.O. & Ashworth, J.R. *Nature* **280**, 116 (1979).
12. McSween, H.Y. Jr *Geochim. cosmochim. Acta* **41**, 1843 (1977).
13. Dodd, R.T. *Earth planet Sci. Lett.* **40**, 71 (1978).
14. Van Schmus, W.R. & Wood, J.A. *Geochim. cosmochim. Acta* **31**, 747 (1967).
15. Leitch, C.A. & Smith, J.V. *Nature* **283**, 60 (1980).
16. Denham, M. Personal communication (1981).

Why is the moon slowing down?

from David W. Hughes

THE Moon orbits the Earth every month and on average moves 13.2° per day across the sky. As long ago as 1695 Edmond Halley (*Phil. trans. R. Soc.* **19**, 1690; 1695) suspected that the Moon was slowing down in its orbit. Now, that the effect has been confirmed, two main causes have been proposed: tidal dissipation and variation in the value of G , the universal 'constant' of gravitation. Measurements of this acceleration have been reviewed and interpreted by F.R. Stephenson of the Department of Geophysics, Liverpool University in a recent edition of the *Journal of the British Astronomical Association* (**91**, 136; 1981).

Tidal friction retards the Earth's rotation. The length of the day is thus slowly increasing at a rate of about one second in 50,000 years. As an autonomous system conserves its angular momentum, the angular momentum lost by the Earth is gained by the Sun and Moon. The Sun is so massive that the reciprocal effect on it is negligible but the Moon does, however, gain a significant amount of angular momentum at the expense of the Earth. This causes our satellite to drift away at a rate of about 4 cm per year.

The acceleration of the Moon can be calculated in two independent ways. The first relies on the analysis of ancient eclipse

observations which extend back over 2,000 years. The difference between the predicted lunar position, if we neglect acceleration, and the observed position which obviously is affected by the acceleration, increases as the square of the elapsed time. Over 2,000 years this difference is as large as two or three degrees. The most recent result obtained using this technique is -30 ± 3 arc sec century⁻². An indirect method based on the observations of the transits of Mercury since 1743 produced a result of -26 ± 2 arc sec century⁻².

The second approach is through the study of oceanic tides. Recent research has shown that the sea level has not varied by more than a few metres over the last three millennia and there is no evidence of any significant change in the lunar tidal friction over the past 2,500 years. It has also been found that the vast open oceans are the major energy sink and not, as previously thought, shallow regions such as the Bering Sea. Calculations of the mean rate of work done by the Moon on the ocean surfaces resulted in an acceleration of -30.6 ± 3.1 arc sec century⁻². Ocean tides perturb the orbits of Earth satellites and an analysis of these perturbations lead to a result of -27.4 ± 3 arc sec century⁻².

The recession rate $\dot{r}$ is related to the accel-

eration $\ddot{r}$ and the rate of change of the constant of gravitation $\dot{G}$ by the following equation:

$$\frac{\dot{r}}{r} = -\frac{2\ddot{r}}{3\dot{r}} + \frac{1}{3} \frac{\dot{G}}{G}$$

If $\ddot{r}$ is measured and $\dot{G}$ is assumed to be zero, $\dot{r}$ can be easily calculated. The present value for $\dot{r}$ is 4.0 ± 0.3 cm yr⁻¹. The Moon was once much closer to the Earth than it is now. G.H. Darwin (*Scientific Papers* **2**, 208, Cambridge, 1908) originally extended this deduction to claim that the Moon may have originated as part of the Earth. The recession rate is inversely proportional to the sixth power of the distance. The assumption that the present rate of energy dissipation has always been active would bring the Moon catastrophically close to Earth less than 1.5×10^9 yr ago. The Earth is three times older than this. Here, to quote W.H. Munk (*Quarterly Journal of the Royal Astronomical Society* **9**, 352; 1968), "we are extrapolating from a few thousand years to a few thousand million years, always a dangerous undertaking, and a harrowing one when the very process is in doubt". Fortunately it has been shown that even moderate changes in the shape of oceans and adjacent seas may materially affect the tidal dissipation and hence the recession rate. The astronomical approach thus seems to be of little help in solving the problem of exactly where the Moon was in the geological past.

The difference between the acceleration measured as a function of Ephemeris Time and the acceleration measured as a function of Atomic Time yields the non-tidal component and consequently the rate of change of G . Some 7,000 observations of lunar occultations of stars have been used by Van Flandern to find an Atomic Time value of the acceleration of -36 ± 5 arc sec century⁻². Laser reflectors have been also placed on the Moon by Apollo astronauts and have been used to give a lunar acceleration with an Atomic Time value of -23.8 ± 4 arc sec century⁻².

Stephenson carefully combines the Ephemeris Time and Atomic Time values of the acceleration. There seems to be considerable overlap between the two data sets and he concludes that the present observational evidence is quite consistent with the interpretation that $\dot{G}$ is zero and the universal constant of gravitation is not varying.

What will be the end result of lunar recession? The day (the Earth's spin period) and the month (the lunar orbital period) are both increasing, the day faster than the month. Eventually it will take Earth fifty-five of our present days to spin once. It will also take the Moon fifty-five days to orbit. Day and month will be equal and the same part of the Earth's surface will always face the Moon. □

David W. Hughes is in the Department of Physics, University of Sheffield.

Intergenic conversion and reiterated genes

from Richard Egel

GENE conversion between reiterated genes is currently gaining recognition as a novel means for exchange in the genome. The term gene conversion stands for the replacement of a particular DNA sequence by a different one. This process is usually initiated by the formation of hybrid DNA between two partially complementary single strands. In general, they belong to two double-stranded molecules. The following discussion, however, deals particularly with the special case when both originate from slightly different parts of the same chromosome.

Until recently, gene conversion has been the esoteric refuge of a few fungal-geneticists who studied the hidden secrets of recombination from aberrant tetrad configurations, joined only by a few determined drosophilists who examined 'half tetrads' of 'attached X' chromosomes, in which the two homologues are connected by a common centromere. These studies suggest that heteroduplex or hybrid DNA is a local intermediate in recombination events of any sort. In a single meiosis, any pair of allelic markers should segregate 2:2 unless they are drawn into a recombination event, in which case various deviations from the regular 2:2 distribution may and do occur. In fungi with four spores per tetrad, 3:1 events are the easiest to detect. In other fungi with eight spores per tetrad, 6:2 and 5:3 events were the first aberrant classes analysed. Originally, such a skewed segregation pattern, where one of the markers had fallen into minority, was classified as 'gene conversion', but it is now common practice to call the formation of hybrid DNA a conversion event, irrespective of whether the final segregation pattern becomes skewed or remains symmetrical.

To understand recombination, one must focus on a single intermediate — the piece of hybrid DNA where two single strands from different molecules have reannealed (Fig. 1). In general, such a reannealing will not lead to a perfect match as any difference in base sequence will reveal itself as a misalignment, comprising single bases or larger loops of noncomplementarity (a). Such discontinuities are either left alone until they are resolved by replication (b), or are erased before replication by repair synthesis (c), using one or the other strand as the conserved template.

If one now considers the adjoining stretches of homoduplex DNA on either side of the hybrid, there are two basic possibilities (Fig. 2): the flanking sequences can originate from the same molecule (a), or they may come from the two molecules that contributed to the

Richard Egel is a professor at the Institute of Genetics, University of Copenhagen.

Hybrid DNA as an intermediate in recombination

Fig. 1 A stretch of hybrid DNA (a), including a noncomplementary region, can be replicated directly (b) or first corrected by repair synthesis (c).

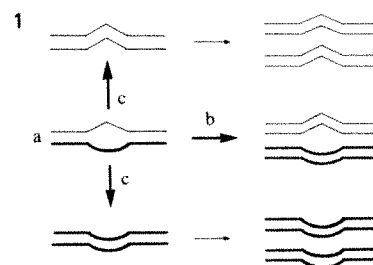


Fig. 2 The hybrid region can be part of a continuous molecule (a) or recombine two different ones (b).

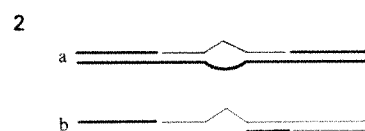
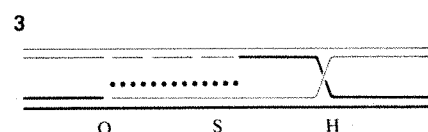


Fig. 3 The 'Aviromore model' of hybrid DNA formation between two molecules¹. Newly synthesized DNA; ———— degraded strand.



hybrid (b). In the first case a limited piece of single-stranded DNA has invaded another duplex without disrupting the overall arrangement of that molecule. In the second, the outer parts have assumed a recombinant configuration.

Usually both the DNA molecules engaging in recombination come off intact. This conservation can be accounted for by the following dynamic model¹, which also takes care of other observations (Fig. 3). At special sites, one of the strands is opened (O), and one of the single-stranded ends begins to invade the partner molecule. The leaving strand is peeled off further, and the gap in its wake is filled by DNA synthesis, using the remaining strand as a template. It may, in fact, be this synthesis that actually drives the reaction. In the receptor molecule, the strand displaced by the invader seems to be degraded initially, but at any moment (S), it can likewise start to invade the partner molecule, whereafter the further formation of hybrid DNA will be symmetrical. This two-phase scenario with a switch from asymmetrical to symmetrical formation of hybrid DNA has been verified experimentally in the fungus *Ascobolus*^{2,3}.

At the final stage (H), the connected molecules are known as the 'Holliday structure' or a 'half chiasma'. The junction is relatively free to move back and forth by the concerted rotation of the four double-stranded arms. Eventually, however, this connected structure needs to be resolved. In meiosis, where recombination is generally thought to occur between two of the four chromatids available after replication, the Holliday structure must be

broken by the cutting of two equivalent strands — the two 'inner' ones or the two 'outer' ones, as placed in the drawing (H). These two alternatives may well assume a less asymmetrical configuration when viewed in three dimensions. In mitotic recombination, however, when hybrid DNA appears to be formed at the two-chromatid stage, resolution is more likely to result from replication through the Holliday structure⁴, rather than by strand cleavage before replication.

A priori, there is no reason why the recombination processes outlined above cannot occur within families of reiterated genes, such as the rRNA, globin or antibody genes. Hybrid DNA can still be formed as long as there is sufficient sequence homology. As the flanking regions may not be homologous, however, their rearrangement by crossing-over will not, in general, be tolerated as this might lead to a change in the arrangement of genes. However, the invading strand can also be disconnected from its source, so there could be recombination between the reiterated coding sequences alone. In this way the overall arrangement of the receptor chromatid would remain the same. This type of information transfer between related but different loci has recently been documented in various organisms.

That recombination essentially depends on local DNA homology is widely accepted. For example, when the *rosy* gene of *Drosophila* is placed inside a heterozygous inversion, the frequency of gene conversion is not changed⁵, whereas crossing-over in the inverted region is

deleterious, generating a dicentric chromosome together with an acentric fragment.

Duplicated genes are often located in a tandem arrangement. This situation can now be simulated experimentally by gene-cloning techniques. Klein and Petes⁶ selected a yeast transformant, in which the wild-type *leu2*⁺ gene, together with a bacterial vector, was integrated in the vicinity of the chromosomal *leu2* gene, which was marked by two frameshift mutations. This strain was then crossed with a *leu2*⁺ partner. Hence, one of the chromosomes of the diploid carried two *leu2* genes in tandem, only separated by the vector DNA. In the absence of recombination, all four spores of a tetrad should receive a copy of the *leu2*⁺ allele and grow without leucine. Occasionally, however, tetrads were observed in which one spore was auxotrophic, due to one or both of the original frameshift mutations. When these aberrant tetrads (12 of 306) were analysed further, the largest fraction (6) was due to conversion between the duplicated genes. Four cases were due to conversion between non-sister chromatids, and only one was due to crossing-over. One tetrad was apparently caused by a multiple event, but it could also be explained by early conversion, if hybrid DNA was formed before the premeiotic S phase — similar to what is usually observed for mitotic recombination. On the basis of these results, the authors conclude that the cells can use gene conversion to maintain sequence homogeneity in a family of reiterated genes. For this purpose, conversion is more efficient than crossing-over at staggered places in a tandem series, the rectification mechanism previously proposed⁷ because with conversion the overall number of genes will be maintained.

Circumstantial evidence for the occurrence of similar events outside the laboratory has recently been detected in the DNA sequences of three human globin genes⁸, two of which were reiterated on the same chromosome with the third allelic to one of the others. As is usual for globin genes, the expressed parts of these genes are interrupted by several introns, and it is one of these that apparently has been subject to intergenic conversion. This conclusion was based on the observation that two equivalent introns of the reiterated genes contained a segment more similar to each other than to the allelic counterpart. Interestingly, this segment was located adjacent to a presumptive hot spot of recombination.

Working with quite a different system, Munz and Leupold⁹ reported genetic evidence for the specific exchange of information between reiterated tRNA genes in fission yeast. They studied the reappearance of suppressor-active alleles that had been inactivated by secondary mutations. This exchange was also most easily interpreted as intergenic conversion.

Intergenic conversion has also been

applied to an analogous phenomenon — mating-type switching in homothallic yeast strains¹⁰⁻¹². It appears that a certain piece of DNA is transferred from a silent position to an expressible site, where it replaces the piece of information previously residing there. This switches the functional mating type to the opposite type. The information transfer is highly asymmetrical in the sense that the source segment is usually left unaffected, and the target segment to be displaced usually discarded. The organism has thus adopted a conversion mechanism for the purpose of gene control, and it has done so by channelling the course into a unique direction.

What other systems are likely to benefit from similar explanations? Gene conversion may perhaps play a role in the generation of antibody diversity. Reciprocal recombination can apparently explain a host of data from V-C joining to class switches. However, the experts still appear divided over whether hypervariable regions can recombine as mini-genes¹³. Their independent incorporation into the framework of the V genes would certainly increase considerably the combinatorial

possibilities. This idea, alas, was badly hit by the discovery of a full set of hypervariable regions in the first V gene from the germ line to be sequenced. However, that blow need not be as fatal as it seems, because only the downright insertion of new sequences is immediately excluded, and not the substitution of one sequence for another of equivalent length. That is precisely what conversion has to offer. So there is still room for mini-genes to jump about, and the search for more evidence need not be called off just yet. □

1. Meselson, M. & Radding, C.M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 358 (1975).
2. Paquette, N. & Rossignol, J.L. *Molec. gen. Genet.* **163**, 313 (1978).
3. Rossignol, J.L. & Haedens, V. *Curr. Genet.* **1**, 185 (1980).
4. Esposito, M.S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4436 (1978).
5. Chovnick, A. *Genetics* **75**, 123 (1973).
6. Klein, H.L. & Petes, T.D. *Nature* **289**, 144 (1981).
7. Petes, T.D. *Cell* **19**, 765 (1980).
8. Slightom, J.L., Blechl, A.L. & Smithies, O. *Cell* **21**, 627 (1980).
9. Munz, P. & Leupold, U. *Alfred Benzon Symp.*, 16 (in the press).
10. Klar, A.J.S., Strathern, J.N., Broach, J.R. & Hicks, J.B. *Nature* **289**, 239 (1981).
11. Klar, A.J.S., McIndoo, J., Strathern, J.N. & Hicks, J.B. *Cell* **22**, 291 (1980).
12. Leupold, U. *Nature, News and Views* **283**, 811 (1980).
13. Robertson, M. *Nature, News and Views* **287**, 390 (1980).



100 years ago

In a paper "presented to both Houses of Parliament" the subject of "*oleomargarine*" as manufactured in the United States is discussed. This substance is made from beef suet by disintegrating in warm water, passing through a fine sieve, melting at 120° F., settling, draining off the oil, and allowing to solidify. During the year ending June 1880, 18,833,330 lbs. of oleomargarine were exported from New York, the greater part going to Holland. The manufacture and sale of this substance is strongly condemned by many butter merchants, and as strongly recommended by various well-known American chemists. Analyses given in the report show very small differences between oleomargarine and natural butter, except in the particular of soluble fats, of which oleomargarine contains considerably less than natural butter.

Cave Animals and Multiple Centres of Species

The readers of Semper's "*Existenzbedingungen der Thiere*," now translated into English, will find (vol.ii, p.268 of the German edition) an interesting discussion of the question of monophyletic or polyphyletic evolution of species, the author decidedly inclining to the latter hypothesis. Considering that at the root of the manifold and difficult problems here involved, there is the relatively simple one of single or multiple centres of each species in a biographical sense, I take leave to ask the following question, hoping for an

answer from among your readers versed in these matters.

To me it seem impossible to maintain the single centres of species in a strict and definite sense without also maintaining the single progenitor of each species, which latter view, formerly considered as a necessary assumption, has been given up by Mr. Darwin in Chapter IV. of the later editions of "*Origin of Species*" (5th ed. p.103, 104). Of course the acceptance of single centres, in the sense of more or less restricted areas of origination, may remain valid for the vast majority of species.

Now, I have sometimes thought that there might be a test for the possibility of multiple centres, which, eventually, would amount almost to an experimental demonstration — namely: *whether there are cases of the same species of blind animals occurring in different caves distant from and without subterranean communication with each other?* Should such cases occur it would be most improbable that the animals in question had been transported from one cave to the other in the modified state, and most probable that they had been independently evolved in each cave from identical species which entered it from without. I formerly noted one instance perhaps in point, viz. a statement of Prof. Cope's (*Nature*, vol.vii, p.11) that "the blind fish of the Wyandotte Cave is the same as that of the Mammoth, the *Amblyopsis speclaeus*, Dekay," but I am not aware whether subterranean communication is, or has been, impossible in this instance. Perhaps more decisive cases have become known of late?

Freiburg im Breisgau, March 4

D. WETTERHAN

from *Nature* **23**, 17 March, 458, 469, 1881.

COMPUTERS IN THE NEW LABORATORY

Where will the revolution end?

Computers have come to stay. And the uses of computing machinery are by no means confined to the first and more obvious applications, the solution of problems in large-scale arithmetic. The pages that follow are intended to point to some of the more conspicuous ways in which computing machinery is being used for the solution of problems in scientific research which are not mere computation. What follows is not, however, a survey of this now vast field but rather a pointer to some of the trends that have become apparent in the past few years. There is every prospect that, in the years ahead, the scientific uses of computing machinery will grow still further. Where will it all end? And what will be the consequences for the practice of scientific research?

One of the most conspicuous developments in the past few years has been the use of computing machinery in the handling of data, but the passage of time serves only to emphasize how much remains to be done. In many circumstances, there is no alternative to computerized handling. It is unthinkable that the data collected by, say, the Einstein X-ray satellite could be handled in any other way. The same is true of the vast amounts of data collected by detection equipment in high-energy physics, some of which yield information in a thousand or so parallel channels in intervals of time no greater than a microsecond. In these respects, computers have literally made possible some kinds of experiments. But data-handling techniques also enhance the value of data. The Einstein satellite is again a striking illustration of what can be done (see page 197). If the detectors (by their nature) record the time at which X-ray photons arrive, or their energy, why not design the data-handling system to exploit this extra information? In this respect, computers can wring more out of an experiment, or a series of observations, than is possible with more traditional techniques of data handling. It is natural that astronomers of all kinds should be alive to the possibilities with which they are now presented. Others will no doubt follow suit.

Another of the important trends of the past few years (hardly referred to in what follows) is the use of computing machinery for simulation. Some of the most ambitious essays of this kind are the climatic models that have been developed at half a dozen centres around the world. The objective is to represent within the circuitry of a computer the physical processes occurring in the real atmosphere. Climatic models are notorious in their demands of the computer hardware, while the development of climatic programs is hampered not only by the complexity of the problems but by ignorance of, or at least uncertainty about, some aspects of the underlying processes. Even so, climatic models have brought valuable understanding to some important aspects of the problem of climatic change. They have surprisingly been less effective at the identification of the physical problems that must be better understood before reliable models can be designed.

Already there have been some striking and successful applications of simulation in the study of the nervous system for example. In other fields, however, simulation by computer model has lived up fully to its purpose, to demonstrate the validity of a theory or to identify its weak points, and to serve as a convenient and rapid means of telling what will happen in the real world. This is another persistent trend.

Computers that run experimental equipment are also on the march. Again, it is unthinkable that the instruments carried aloft by rockets and satellites could be operated differently, but the sophistication of the technology now in use is a quite remarkable demonstration of what has been accomplished in the past twenty

years, partly no doubt because of military interest in the same techniques. That experimental equipment in, for example, high-energy physics should be controlled in the same way is also easily understood — there is no choice. With the introduction of relatively cheap microprocessing equipment, however, other surprising applications of this kind have come to light. In some quite small laboratories, it turns out to be worthwhile to use such a piece of machinery for controlling a piece of standard equipment from which a sequence of observations is required. The benefits are certainty, speed and the saving of research assistants' time. The cost is often small.

In the nature of things, however, most excitement in the years ahead will lie in directions not now foreseen. To many people, the use of computers for carrying out routine algebraic calculations (see page 198) will be a surprise. With hindsight, of course, this development seems natural enough. But how many would have staked money or reputation on the feasibility and the success of this development ten years ago, when computer algebra was in its infancy? And who now can guess what other developments, perhaps more radical, are buried in embryonic form in the research programmes of the computer laboratories?

Many of the applications of computers in past decade have made possible tasks that otherwise could not (or would not) have been attempted. Most of the same applications have been conceived of by the people most concerned to benefit from them — researchers themselves, not the manufacturers of computer hardware. And the chances are that the burden of innovation will continue to fall on the ultimate beneficiaries. With this in prospect, there would seem to be an urgent need that researchers in all fields, not simply those in the physical sciences, should be equipped by their education to be able to appreciate the possible benefit of computer applications. Although, as the experience of recent decades suggests, the most unlikely disciplines appear to accrete their own bands of computer enthusiasts, it seems unwise that they should rely on such a random process. The implications for the education of graduate students are obvious. Some practical experience of how to put a computer to good use is essential. But it may be even more important that graduate students should be led to understand the limitations of computer machinery — and to appreciate the pitfalls of over-dependence.

Like all blessings, the new computer technology is, after all, a mixed blessing. Making computers work well can rarely (outside the computer laboratories) be an end in itself. And the pace of discovery would be constrained if too many working scientists spent too much time writing programs for machines. It may also be damaged if the seductiveness of computing machinery, and the temptation to ask too much from simple computer programs, leads laboratories into spending too great a proportion of their resources on what is, after all, merely a means to an end. □

CONTENTS

Computers and the biologist	194
Computers cheaper: Computers in physics	195
Logic circuits before computers	196
Astronomy by numbers	197
Some things old, some new	197
Algebra made mechanical	198
Plenty of programs: Algebraic computer programs	200

Computers and the biologist

Among biologists, physiologists have inevitably taken most readily to the use of computers; people brought up to know the difference between d.c. and a.c. amplifiers are naturally more inclined to give house-room in their laboratories to further electronic equipment. Yet the past few years have seen the infiltration by computers of most other kinds of research laboratories.

In all kinds of laboratories, the chief use of computers is still in the storage and analysis of data. The principles are as old as computers themselves. Data are digitized, stored within the accessible part of the computer or dumped onto magnetic tape, and analysed by whatever computer program is appropriate. In the 1960s, most large biological laboratories found it expedient to install substantial main-frame computers to provide their members with a data-processing service. Now, however, with the arrival of mini-computers and microprocessors, the trend is towards distributed computer capacity. Individual laboratories are acquiring their own mini-computers, while many pieces of equipment are now designed to operate with their own built-in microprocessors. (Most manufacturers of substantial pieces of equipment, for radioimmunoassay, for example, now provide floppy disks which allow the users to play games such as *Star Wars* as well as the more sober programs required to analyse the data, no doubt from compassion for those working in laboratories with time on their hands.)

This trend is probably inevitable, and unstoppable. For the user, the obvious advantage is to be fully independent, able to process one's data whenever necessary. Central machinery working on a batch-processing basis is necessarily by comparison an encumbrance. And there has grown up a certain air of disenchantment, at least among biologists and biochemists, with computer systems designed to be accessible by means of computer terminals in the laboratory.

Nevertheless, some of those concerned with the provision of computing facilities in general purpose laboratories are uneasy. Although the new small computers and microprocessors are cheap, the software required for specific specialized function may make serious demands of program-

ming skills. The end result, some gloomy calculations suggest, may be that the central cost to laboratories may be greater than at present. The alternative, that individual researchers should themselves become skilled at programming their own machines, is regarded with apprehension.

Laboratory administrators are also uneasy about the cost of measuring equipment with built-in microprocessors. The convenience and efficiency of these machines is beyond dispute. The suspicion is merely that the value added by the manufacturers of the equipment may not give the user a fair share in the cost benefits of the cheaper hardware now available.

The truth is probably that laboratory administrators are learning a painful lesson in economics. The development cost of software for a commercial measuring instrument is likely to be greater than that of writing a program in a particular laboratory. But the manufacturers of the measuring instruments with built-in microprocessors also find that demand is buoyant, suggesting that people able to afford the new equipment are prepared to pay for its convenience.

There are several well-attested illustrations of the costs that can be saved by home-made equipment. Thus one West Coast neurophysiologist, unable to afford from his current grant the \$20,000 to buy a suitable wave-form analyser, reckons that he was able to produce a better system for a quarter of the cost and at the same time produce a system of analysis better suited to his needs. The underlying dilemma is familiar — to what extent should research laboratories spend their time and energy on the construction of instruments which are nevertheless obtainable commercially.

Not all of the tasks of the general biological laboratory have as yet been delegated to the instrument manufacturers, however. Indeed, even in some of the most familiar tasks, the analysis of electrical output from physiological preparations for example, formidable intellectual difficulties keep cropping up.

The standard record in such experiments consists of a voltage (or several voltages) varying with time, on a time scale of milliseconds or thereabouts. Given a visual trace of such a voltage, it is possible with a ruler to estimate parameters such as the peak voltage, the time taken to reach that voltage and so on. Computer systems store the same information in digital form, and there are constant (if groundless) worries about the extent to which such data embodies "everything". More seriously, physiologists are likely to insist that some form of visual record should be obtained in parallel with whatever computerized analysis is found convenient. Machines may be able to analyse complicated sets of data once they have been provided with criteria for setting about this task, but

cannot recognize novel kinds of patterns.

So much is evident in the difficulties which keep cropping up in the research programmes intended to provide more sophisticated monitoring of the physiological condition of hospital patients, people in intensive care units for example. Many of the systems developed for this purpose depend on the recognition of abnormal trends in, for example, heart-beat rate, or variation from one cycle to another. One difficulty is that too little is known about the statistical properties of normal heart-beats to provide a sure yardstick for abnormality. The result is that false warnings abound, while few are confident that the systems now in use are capable of recognizing more than a small proportion of impending crises.

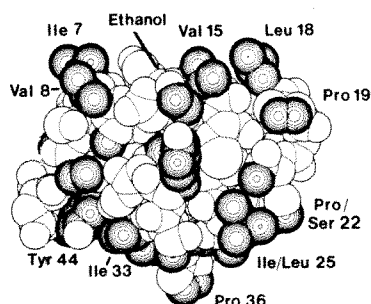
But what of the application of computers to tasks in biological research that have hitherto been beyond the scope of conventional techniques?

Many of the innovations in the past few years are comparable with those in other fields. It is unthinkable that X-ray diffraction data should, these days, be analysed by the traditional methods, and computer techniques are used as a matter of routine in construction of the models of most molecules, biological or otherwise. Quite apart from the way in which computing machinery has given people the courage to tackle problems that would otherwise have been neglected, the more economical use of data has made it possible to work with lower X-ray intensities.

Another development of this kind is in the construction of molecular models. Modelling biological molecules of any complexity is a formidable task, and this is the principal incentive for the development in the past decade or so of computer programs which can handle such tasks quickly but, more important, can in principle explore all possible configurations of a complicated molecule.

These are essentially data-handling tasks. The most recent development of this kind has been prompted by the sheer size of the nucleotide sequences now being accumulated. It is natural that these should be stored in computer form, and plans for the development of such a computer bank have been discussed between the National Institutes of Health in the United States and the European Molecular Biology Organization. One obvious elaboration of such a system is the use of computer programs for searching nucleotide sequences for particular configurations of nucleotides, for example for configurations that define the points of action of particular restriction enzymes. But the construction of a complete nucleotide sequence is not always straightforward, given that it may have to be inferred from the nucleotide sequences of overlapping fragments of a complete DNA molecule, which is why there are now ambitions to solve such essentially combinatorial problems by using computing machinery. □

Computer model of the protein crambin



Computers cheaper

One of the most surprising additions to the language brought about by the use of computers in high-energy physics is the verb "to Monte Carlo". (It is not clear whether the root of the verb should be hyphenated.) The verb is a reference to the now standard procedure by which somebody planning a high-energy physics experiment must recognize that he will not get the grant he needs unless he has first simulated the efficiency of his detector set-up by means of a computer calculation. The principle is that the path of any particle through the sequence of detectors and other recording devices that may be involved in an experiment will not only depend on the origin and momentum of the particle within a beam (whose dimensions are never negligible) but also on the strictly statistical chance that it will interact or decay at some stage. Analytical calculations are rarely possible; instead physicists use Monte Carlo calculations, in which the hypothetical paths of individual particles with parameters chosen by means of random number generators, and chances of interaction similarly determined, are followed through the equipment in the computer simulation. Successful applicants for grants and successful experimenters know that much depends on the skill with which their experiments have been "Monte Carloed" in advance.

This, however, is a relatively indirect use of computers in the high-energy physics laboratories. A much more direct use of computing machinery is in the control and management of experiments, and here perhaps the most striking application of computer techniques is in the control of the beam of 400 GeV protons in the SPS proton synchrotron at CERN.

The CERN system

The problem is easily stated. The machine consists of a vacuum chamber arranged roughly as a spherical ring 2,000 metres in diameter. The circulating beam of protons is intended to have a fixed orbit within the vacuum chamber while, to prevent collisions with the walls of the vacuum chamber and thus the loss of protons, the diameter of the beam itself should be of the order of 1 centimetre. But the beam circulates in the form of a pulse or package of protons, thereby increasing the electrostatic repulsion between them and adding to the inevitable space charge whose consequence is further to spread out the proton beam.

To avoid this difficulty, the proton synchrotron has been provided with a control system consisting of magnets for correcting the position of the beam within the vacuum chamber based on signals picked up by detectors elsewhere in the ring. It is tempting to describe this as a way in which the position of the beam is

adjusted in the light of information about its characteristics a few microseconds earlier, at the other end of the 2,000-metre diameter ring. In practice, of course, that description is an over-dramatization. It is entirely possible to accommodate departures of the proton beam from its intended position, or its spread beyond the intended width, for several orbits around the ring.

Computer control

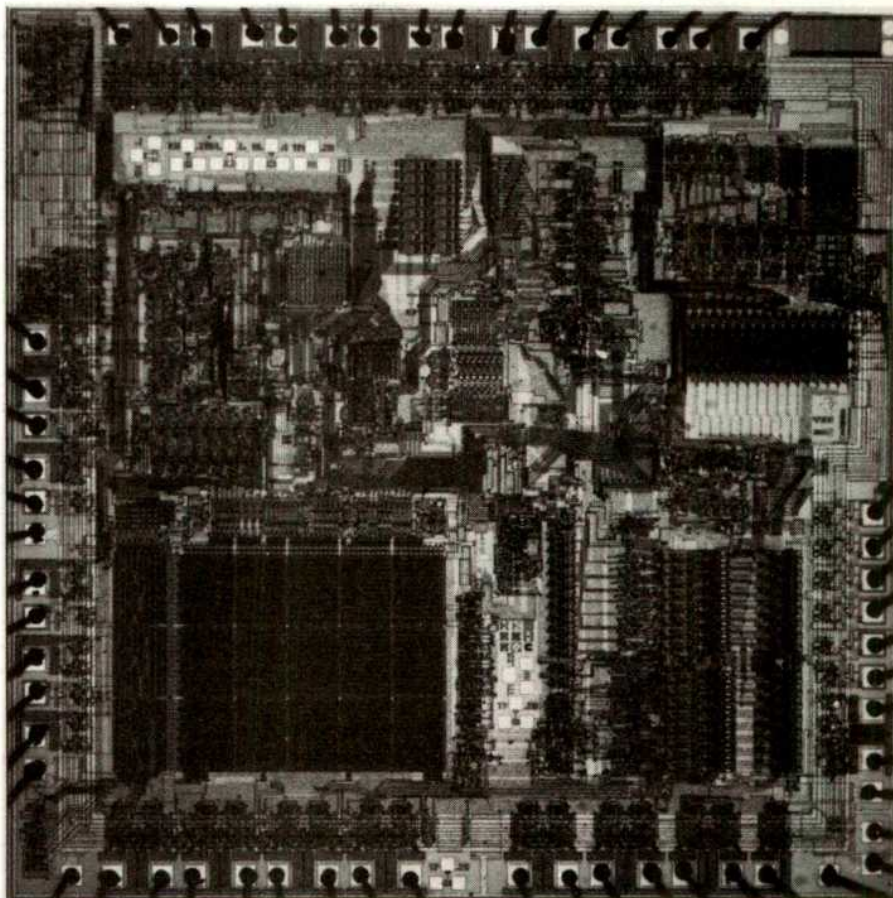
Inevitably the beam adjustment has to be controlled by a computer. While in principle the system is not markedly different from those which make decisions on the ways in which detection equipment should operate based on observations of the properties of a particular particle entering the equipment, there is a sense in which the development of these systems has helped to make very high-energy accelerators feasible — or at least to reduce their cost or to improve their efficiency. The quality of a beam, after all, depends on its intensity (the number of particles in a single pulse) and on its small dimension. Too many particles lost in collision with the walls means too much beam intensity sacrificed. But the obvious solution to this problem, to use a larger vacuum chamber or stronger focusing magnets, adds

enormously to the cost of what is, after all, the most expensive part of such a machine, the magnet system in the poles of which the vacuum chamber sits.

Inevitably, such sophisticated control systems, already in use at CERN and at the Fermilab in the United States, will be incorporated in all large machines built in the years ahead. When funds are hard to come by, and when high-energy physicists must necessarily be seen by their sponsors to be making all possible economies in their proposals, it is clearly advisable to spend a little money on the computer control of the particle accelerator so as to save what may be a substantial percentage of the total cost of electromechanical construction.

Bubble chambers

These uses of computing machinery in high-energy physics are inconspicuous although now familiar. During the past 20 years, however, one of the most eagerly sought developments has been that of the computerized handling of the photographs of the tracks left by charged particles in bubble chambers — the modern versions of the Wilson cloud chamber. Tracks in bubble chambers consist literally of strings of bubbles forming in, say, liquid hydrogen at points where electrical ionization has been created. Especially with the improvement of accelerating machines and the inexorable trend towards higher energy, bubble chamber photographs (which come in stereoscopic pairs or



A sixteen-bit microprocessor — many thousands of transistors on a ¼ square inch of silicon

Texas Instruments Ltd

...multiples) consist mostly of uninteresting tracks left by particles which have passed through the chamber unaffected (although perhaps deflected in one direction or the other, according to their charge, by the external magnetic field of a magnet). Traditionally, the analysis of such photographs, like the analysis of tracks in photographic emulsions, has been carried out in high-energy physics laboratories by small armies of part-time workers, usually women. Pressure to computerize the task of analysis has arisen not merely for economic but also for managerial reasons. A complete analysis of all the interactions recorded on a single bubble chamber image is too great to be attempted except in exceptional circumstances. The first attempts at simplifying the task required that the human operator should follow only tracks judged to be interesting with some device whose position in two directions could be recorded automatically, by pressing a button perhaps, stored in digital form and analysed to give a 3-dimensional reconstruction of the track.

In the past 15 years, however, the analysis of the complicated sets of images of which only some are interesting has been further simplified by the development of the so-called Hough-Powell device, essentially a means of scanning a plate or of scanning an image photometrically with a flying laser spot constrained on a predetermined raster. A digital output is thus provided in a simple form — each spot on the image is either black or white.

Limitations

In principle, appropriate analysis of two stereoscopic photographs of this kind can reconstruct the tracks of all charged particles passing through a bubble chamber. In practice, even with the use of computer hardware, the task is too slow and too demanding of storage space. Accordingly, it is still standard procedure for somebody, by means of visual inspection, to select those particular tracks in an image which are interesting enough to justify detailed analysis. One by-product of these computer systems is that they yield not merely a detailed geometrical reconstruction of a track but also a kinematical analysis of those particles involved. Whether they will ever be capable of operating entirely automatically, ignoring uninteresting tracks and singling out those which, by predefined criteria, are especially interesting or significant is another matter — and a problem in pattern recognition.

The development of these systems has been to an encouraging extent a collaborative business, in which the large high-energy accelerator laboratories have played a central part. In Europe, for example, one direct result of this collaboration has been the development of a set of standards (called CAMAC) intended to ensure compatibility of information processing equipment. □

Logic circuits before computers

When nobody was especially well supplied with equipment, but when nuclear physicists were better supplied than anybody — in the late 1940s and early 1950s — it was relatively easy to predict who the first users of computers would be. It is tempting to think, at least where the nuclear physicists are concerned, that they learned their habits during the Manhattan Project. In truth, the search for ways of applying computational techniques to the task of observing what you want to observe and not some tedious background signal came two decades earlier, in 1931.

Then, for practical purposes, there were merely a handful of techniques for observing the effects of energetic charged particles, one of which consisted simply of measuring the rate at which a charged electroscope would lose its charge. The most interesting technique was C.T.R. Wilson's cloud chamber, which had for the first time made it possible to visualize the tracks of charged particles travelling through its confined volume, but which to begin with could take snapshots of the flux of charged particles through it only at randomly chosen intervals. The chance of finding something interesting cannot have been much greater than that of finding a needle in a haystack.

Geiger counters, also widely used in the 1920s, were more discriminating than cloud chambers in that they would produce an electrical discharge only if the energy of the charged particle traversing them was greater than a certain amount determined by the quantity of precise composition of the gas with which the vacuum was filled — so much argon, so much nitrogen and so on. The snag with a Geiger counter (1930s vintage) is that it was simply a device for answering "yes" or "no" to the question "Where (and when) did a charged particle with more than the threshold energy pass by?"

But what more natural than to couple a layer of Geiger counters connected in parallel to the expansion mechanism of a cloud chamber so as to ensure that the cloud chamber photographs would include the track of at least one charged particle whose energy was greater than the threshold energy? This, in 1930, is what P.M.S. Blackett and G. Occhialini (then at the Cavendish Laboratory) set out to do. Blackett used ruefully to complain that if Occhialini had not taken a long holiday, they would have found the positron before P.W. Anderson.

From that point on, the use of elementary logic circuits, answering "yes" or "no" to questions such as "Has such and such a kind of charged particle passed through?", was inevitable. Arrays of Geiger counters working in coincidence ("A charged particle went in and another came out the other side") or

in anti-coincidence ("It went in, but nothing came out") were all the fashion. Chadwick proved the existence of the neutron by such a technique.

By the outbreak of the Second World War, people's ambitions had grown, while detection techniques had become more complicated — sophisticated is hardly the word. By encasing cloud chambers in magnetic fields, and arranging that the chamber would not be triggered into action unless a charged particle "went in" and emerged virtually unchanged in direction (indicating that its energy was very large), it was possible to ensure that all cloud chamber photographs included at least one track of an energetic particle (which is how the "strange" particles of matter were eventually discovered, in 1948).

Long before then, however, logic circuits and the development of detecting devices had made it possible not merely to



P.M.S. Blackett

record the arrival of a charged particle but to assign it to some range of energy. The outputs from the devices used for such purposes were essentially digital, so why should they not be accumulated in digital form? This is how it came about that nuclear physics laboratories in the 1950s, but especially those connected with semi-secret enterprises, were the first to be blessed not merely with huge racks of electronics fitted out with tube-driven logic circuits but with slave electric typewriters as well, recording the data the detectors generated. (Most secretaries in those same laboratories still used manual machines.)

Frivolity apart, the outcome has been important. The high-energy physicists have been among the leaders in the use of computers for the control of experimental equipment. Most of the detection equipment now run in connection with high-energy particle accelerators consists of an elaboration of the counter-controlled cloud chamber now exactly half a century old. The difference is that the primitive logic circuits of the old devices have been replaced by computing machinery. □

Astronomy by numbers

Computers have taken astronomy by storm in the past decade or so, and for obvious reasons. Need has been matched by opportunity, but astronomers have also been temperamentally attuned to computerization by the space programmes of recent years. Once it became necessary to work with data transmitted in digital form by telemetry from rockets and satellites, it was certain sooner or later to become habitual to use the same techniques for dealing with other sources of data.

Astronomy has traditionally been profligate in its regard for data. A single plate from a large Schmidt camera, covering a patch of sky just six degrees in each dimension, may include up to 10^6 images of different kinds. Merely recording the positions of these images would occupy a skilled technician for the best part of a lifetime. The libraries of the world's observatories most probably contain enough information to keep the world's astronomers usefully employed for the rest of their lives — perhaps on problems of no particular interest.

The simple solution to the problem of counting stars and measuring plates is similar to that to which high-energy physicists were forced in the 1960s, when it became impracticable to measure the tracks of fast particles in bubble chambers

by hand and eye. The first need is a device for converting star images into numerical records of the density of exposure at different points on a plate. Plate reading systems have been developed at several observatories. As always, the tricks lie not in the hardware but the software — in this case the rudimentary programs for telling which image is what. Is it a star? Or a galaxy (and if so which kind)? Or something else?

One neat way of describing the benefits of these systems is that it has taken some fifty man-years of programming time to develop the software used with the COSMOS plate measuring system at the Royal Observatory, Edinburgh and that the result is to make possible the measurement of the images on a single Schmidt plate in one day, not one lifetime (also fifty man-years). This figure of merit is also an index of the true cost of these devices — £500,000 is probably an underestimate of the true cost of developing COSMOS.

The other component in the recent computerization of astronomy is the digitization of the data base. The most striking illustration so far of the benefits of this technique is the Einstein X-ray satellite (now, happily and surprisingly, restored to working order). The essence of the X-ray telescope is a system of conically shaped

glass surfaces from which X-ray photons are deflected at more or less grazing incidence. The detectors are inherently electronic, borrowed from nuclear physics. So the data are inherently digital in form, literally consisting of the spatial coordinates in the image plane of individual X-ray photons. So why not record not merely the angle from which these photons have arrived but the time of their arrival and even information about the energy of the incident photons?

In practice, the Einstein satellite is equipped with a variety of X-ray detectors which separately provide different kinds of data. The detector with the highest resolution, called the High Resolution Imager, is designed to record the arrival of X-ray photons on the MgF_2 plates from which cascades of electrons are generated to within 1 second of arc and within 8 milliseconds. The working spatial resolution of the instrument is less than this (some tens of seconds of arc) because of the properties of the mirror system, but this high resolution instrument provides no spectral information worth speaking of, with the result that the computer systems store merely spatial and temporal information.

Other instruments in the Einstein Observatory, which can generate spectral information (but whose spatial resolution is not as high), yield data streams from which spatial, temporal and spectral information can be regenerated. The

Some things old, some new

Electronic computers are merely the latest stage in the computerization of the laboratory. The earliest aids to laboratory computation were probably the astrolabes and armillary circles going back to the antiquity of astronomy. The first explicit

aids to laboratory computation were probably Napier's logarithmic systems of coloured rods (called "bones") introduced early in the seventeenth century, and quickly followed by Pascal's logarithmic scale, much used (with the help of a pair of dividers) by navigators in the eighteenth century.

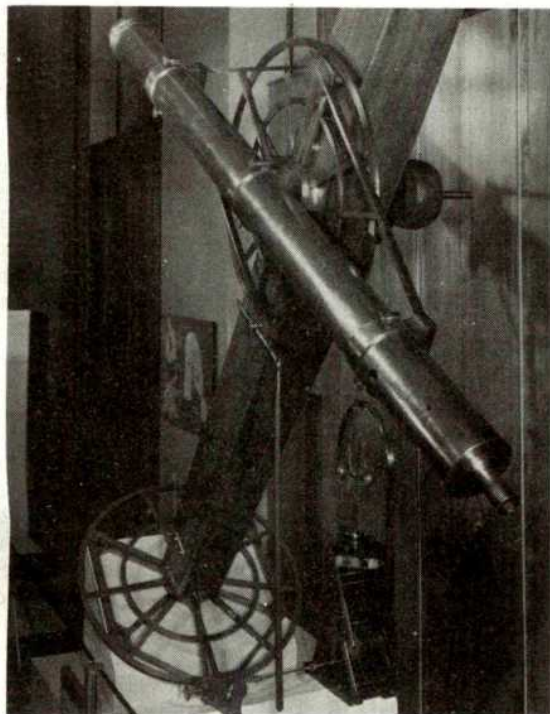
The needs of astronomers stimulated interest in what may have been the first computerized aid to laboratory observation — sky-following telescopes. Robert Hooke appears to have been the first to have conceived of this innovation, but his proposal for a clockwork-driven telescope was built for G.D. Cassini at the Paris Observatory and commissioned in 1682. This was a huge device, with the objective mounted at the centre of a clockwork-driven structure and with the

eyepiece mounted separately at a distance of more than 100 feet. Cassini's telescope was known as his *machine parallatique*.

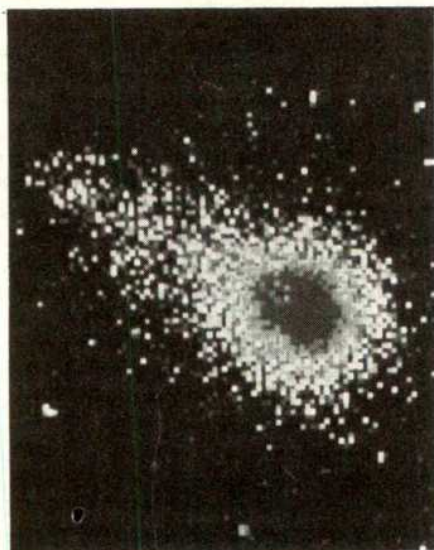
Attempts at the further development of clockwork-driven telescopes appear to have been few and far between in the eighteenth century. Such instruments seem to have come into their own only in the early decades of the nineteenth century.

Dr John Darius (Science Museum, London) writes:

Driving telescopes around their polar axis to follow the diurnal apparent motion of the stars is useful for visual observations but absolutely vital for astronomical photography. The earliest drives, dating from the beginning of the nineteenth century, were purely mechanical, depending on a clockwork mechanism driven by weights to rotate the telescope in the equatorial plane at the rate of 15 degrees of arc every hour. By around 1870, telescope drives regulated by an electric pendulum clock had been invented, but by 1910 electric motors were coming into regular use. Nowadays digitally controlled synchronous motors as uniform as the frequency of the mains supply have almost wholly replaced their historic forebears. The most modern systems now use servomechanisms which are controlled by computer. □



W.H. Smyth's clockwork-driven telescope, 1829



An image of comet Bradfield taken by the Fine Image Sensor of the IUE satellite observatory

practical value of this complication is that some X-ray sources are especially interesting because of their rapid fluctuation with time. Thus some X-ray sources that consist of binary star systems in which material is being transferred from one star to the other intermittently fluctuate in X-ray intensity on time-scales of a few minutes. Others, for example quasars, also fluctuate in time. Some of the earliest observations with the Einstein Observatory (see, for example, Tananbaum, H. *et al.*, *Astrophys. J. Lett.* **234**, L9; 1979) were used to demonstrate variability of X-ray intensity from distant quasars over a period of a few hours on the basis of a mere 50 or so X-ray photons.

It goes without saying that the planning of a series of observations with such a satellite requires that the on-board computers should be explicitly instructed on what information to store, and in which form, while the planning of the way in which a high-resolution instrument scans a small patch of sky requires some forethought about the way in which the data will be analysed. Programs by which the field of view moves in overlapping steps across the sky are often advantageous. It is commonplace that observations of particular objects are pre-programmed not for arbitrary periods of time but in such a way as to yield statistically significant results (which entails concurrent measurement of the background radiation).

Astronomy as a whole now seems destined to move in these directions, helped by the development of a variety of detection devices whose chief purpose has been to improve the sensitivity of telescopes (optical, radio, infrared, what-have-you) but which as a by-product yield a stream of numbers as the crude data. Quite what balance will in the end be struck between positional, spectral and temporal information is, however, still very much a matter for astronomers to decide. For the plain truth is that while it may in principle

be possible to record every detail of every quantum event occurring at the light-sensitive surfaces of the image intensifiers, there is no hope of being able to store such a wealth of data.

Accordingly, there is no sense in which the computerization of astronomy can be said to have robbed the craft of the need for discrimination and judgement. Although the coming of more efficient light detectors has often made possible dramatic reductions in the length of time needed to gather significant information about some distant object, observing time is still at a premium. The need for deciding clearly in advance what questions it is hoped to answer, and for collecting the smallest amount of data required, has therefore not diminished. Indeed, it has been intensified by the ease of discrimination in the collection of data.

The bugbear that remains is that the craft has become almost inextricably bound up with computer programming. Research students have repeatedly found themselves caught up in projects for writing programs to recognize, say, elliptical galaxies, or to measure the ellipticity of globular clusters. Educative though these exploits may be, there is obviously a danger that too much creativity may be lavished on what is essentially repetitive work.

This is part of the incentive for the development in the United Kingdom of a linked network of six astronomy centres in a project funded by the Science Research Council under the trekky name of STARLINK. The plan is that five British centres (the Royal Greenwich Observatory, the Royal Observatory Edinburgh, University College London and the Universities of Manchester and Cambridge) should be provided with identical computer systems and linked by low-capacity data links to a coordinating centre at the council's Rutherford-Appleton Laboratory.

The objective is to make possible the more economical use of people's pro-

gramming energies. Schemes for the development of novel programs will be worked out on an *ad hoc* basis by the collaborating laboratories. Programs but also images (containing 500×500 picture elements or pixels to begin with, perhaps 16 times as many later on) can be transmitted from one centre to another. There are some ambitions that the system may be extended internationally, to Europe, possibly the United States and, with luck, to Australia.

The STARLINK system is interactive in that the DEC VAX computer system provides astronomers with a graphic display of an image which can be manipulated this way or that, to yield profiles across an object for example. Much of the work so far, however, has been the development of programs for the analysis of spectra. How does one extract from a record of intensity as a function of wavelength numerical data for line frequencies and line depths? Superficially, the problem resembles that which crops up in quite different fields. In medical physics, people are forever trying to decide how best to characterize heartbeats.

The system is not yet fully a going concern, although it is hoped that a full package of software will be available "in months rather than years". For some purposes, however, the response of the system is plainly not fast enough — people who want to follow the motion of features in the atmosphere of Jupiter, for example, need more sophisticated image processing, with a much faster response. This is why, for instance, some of those concerned with the processing of data from the recent series of planetary probes have independently developed more sophisticated image processing systems. One such is the IPSIS system at University College London, which gives scientists access to a large number of satellite images stored in digital form. The interactive nature of the system permits a detailed examination of all aspects of the data. □

Algebra made mechanical

The notion that computers are mere number crunchers is fast being exorcised by one small group of people — the computer algebraists. In the past twenty years, several classes of problems which require analytical and not numerical solutions have been tackled successfully. Now, for example, there are computer programs that can yield analytic expressions for indefinite integrals a good deal more reliably than the "random undergraduate", as one practitioner puts it.

Neither the need nor the feasibility of these tasks is self-evident. And, as things are, the community of computer algebraists is small and scattered. Yet, in some fields, it is now commonplace for people with time-consuming algebraic calculations on their hands to use a computer program for the purpose. And

the fields in which these techniques are being used are so diverse that there seems every prospect that computer algebra will spread.

The feasibility of computer algebra is demonstrated easily enough — perhaps most easily by the manipulation of polynomial expressions such as $a + bx + cx^2$. Adding second-degree polynomials together requires an algorithm for adding together the coefficients. Multiplication and subtraction are similarly straightforward. Division is (as always) more difficult, but there is no reason why the rules of school algebra should not again be applicable.

Producing analytical forms of indefinite integrals, or analytic solutions to differential equations, is necessarily more difficult, not merely because the computer

programs used for these purposes must in some sense "know" the essential mathematical properties of functions such as $\sin x$ or e^x , but also because the analogy with the way in which students are taught to solve such problems (by classification around solvable standard forms) is not sufficiently general.

Yet programs for these purposes have been developed at several centres. Put to work on sets of problems taken from standard textbooks, they typically yield success rates of 97 or 98 per cent — they have more than the competence expected of undergraduates but are not yet good enough to satisfy the purists. But these, of course, are not the fields in which computer algebra is likely to make its mark.

There are many more telling applications of the technique, in fields where in the past heroic individuals have battled with forbidding mountains of algebra or have, as likely as not, decided that discretion is the better part of valour.

The successes

One of the first and most striking successes in this field was in celestial mechanics — the theory of the Moon. The problem is straightforward but daunting. Since Newton's time, a knowledge of the geocentric longitude of the Moon as a function of time has been an essential ingredient of the craft of navigation. To state the problem is simple enough — it is a classical problem in Newtonian mechanics. It is, however, a three-body problem even without the gravitational effects of Jupiter and the other planets and is thus incapable of yielding an exact solution.

In the circumstances, only perturbation methods of solution will suffice. Numerical solutions, widely used, have one serious and for some purposes insuperable drawback — errors accumulate horrendously as the span of the time variable increases, so that it is not, for example, possible to be sure of where and when some past eclipse should have occurred. Accordingly, there has always been great interest in analytical solutions of the problem of the Moon — essentially approximations in the form of convergent series whose terms are individually analytical.

The only successful attack on the problem "by hand", so to speak, was that of Delaunay, whose two huge volumes on the subject were published by the French Academy in 1860 and 1867. The first hundred pages of the first volume are entirely given over to a representation of Delaunay's analytical expression of his "disturbing function" (essentially the departure of the gravitational potential of the Moon from simple Newtonian calculations of the potential due to the Earth). This seems to have represented literally a lifetime's algebra — Delaunay carried through his calculations to the seventh order in small quantities. Now, the same

calculation can be run on a computer in a few tens of seconds, and to a higher order of approximation. Delaunay's descendants need not be ashamed that the first computer programs showed that his algebra had not been immaculate.

Similar analytical solutions of more intricate (and previously intractable) problems in celestial mechanics are now matters of routine — for example, the movement of putative Earth satellites about the Lagrangian points at which the gravitational forces of the Earth, Sun and Moon cancel out. The practical importance of these developments is considerable, if only because the error of an analytical approximation is more easily assessed than that of a numerical approximation.

General relativity

Curiously, as it may seem, one of the other well tried (and much used) applications of computer algebra is in general relativity the problem is this. Einstein's general theory of relativity is general in the sense that it is possible to derive the gravitational field equations, typified, for example, by the gravitational forces on a test particle itself of negligible mass from an analytical expression for the metric of space-time. The simplest possible form of the metric is

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 t^2$$

where c is the velocity of light, x , y and z are cartesian coordinates and t is time. Substituting this metric into Einstein's general theory merely recaptures the kinematic aspects of special relativity, as of course it should.

More generally, the metric of space-time is written

$$ds^2 = G_{\mu\nu} dx^\mu dx^\nu$$

where the indices μ and ν run from 1 to 4, x^μ is one of the four coordinates of space-time and the quantities $G_{\mu\nu}$ are, in general, functions of these coordinates. The underlying convention adopted by the relativists is that indices which occur twice represent summations — in other words, the general metric consists of 16 terms of which the first is

$$G_{11} dx^1 dx^1 = G_{11} (dx^1)^2$$

Regarded as a general theory, the General Theory is intractable. The field equations, the differential equations which predict how particles should behave, are nonlinear and therefore soluble only with special assumptions. So the standard way of tackling problems in general relativity consists either of guessing that a metric might illustrate some particular property of a gravitating universe, and to calculate from that the field equations, or alternatively to look for approximate solutions of the field equations, valid in some narrow regions of space and time. In either case, the hapless relativist finds it necessary to calculate what are called the Christoffel symbols of which in general there are 64, each of them a combination of three first partial derivatives of the functions which define the metric. That, however, is only a

beginning. The next step is to calculate the Riemann-Christoffel tensor, which is in general a tensor with 256 distinct components, each of which is a linear combination of partial differential coefficients of the Christoffel symbols.

The sheer algebraic complexity (and tedium) of these calculations explains why only a few relatively simple metrics have been used in general relativity. Even the simplest of the more general metrics, the general case of an orthogonal metric

$$ds^2 = D(dx^4)^2 - A(dx^1)^2 - B(dx^2)^2 - C(dx^3)^2$$

in which the symbols D , A , B and C stand for functions of the four coordinates x^i etc. presents formidable problems of computation. Herbert Dingle in 1933 published a list of the Christoffel symbols (*Proc. natn. Acad. Sci. U.S.A.* **19**, 559; 1933) which represented the fruit of several years work and which remains a standard reference list. The likelihood that such calculations would now be attempted with more general metrics is exceedingly small, which is why computer programs for making these calculations are widely used.

Probably the most widely used set of computer programs with applications in theoretical physics is, however, that called SCHOONSCHIP, developed in the late 1960s at CERN in Geneva as a means of calculating physical quantities in quantum electrodynamics. Briefly, such problems invariably create the need to calculate the trace (the sum of the diagonal elements) of a four-dimensional matrix which is itself the sum of what may be a large number of four-dimensional matrices obtained by multiplying together the so-called Dirac matrices. Although the elements of the Dirac matrices are simple (either 0 or 1), the algebraic task of computing by hand the product of say eight of them is enormous, and the chance of making a mistake is very great. (In practice, the multiplication of Dirac matrices goes hand in hand with the evaluation of four-dimensional integrals, now also computerized.) In some applications, as for example in the application of similar techniques to unified field theories or quantum gravity, many vital calculations have not been carried out simply because of the complexity of the task. Inevitably, with the passage of time, programs such as SCHOONSCHIP have therefore become widely used.

Versatility

These three applications in theoretical physics of computer programs are tangible and valuable. But if the object of computerizing algebra is to provide theoretical physicists with what Professor Joel Moses has called an algebraic assistant, surely computer systems should be more versatile. This consideration underlies the attempts which have been made since the late 1960s to develop general-purpose algebraic computer programs, the best known of which is called MACSYMA

and lives in a PDP-10 computer at the Massachusetts Institute of Technology (MIT). A similar computer system called REDUCE was developed during the same period at the University of Utah.

In principle, these programs are designed to deal with whatever algebraic problems may be thrown at them. In their development, handling polynomial algebra and matrix multiplication was relatively straightforward. The problem of integration, requiring that the program should make decisions about the most appropriate way of solving a problem, was necessarily more difficult. Even differentiation creates problems when irrational functions are involved. Now, however, MACSYMA has more than a thousand users, many of whom make use of it by means of international communications systems such as the ARPA network. Versions of the program are soon to be made available in a dialect of the computer language LISP, when MACSYMA is likely to be as easily available as its competitor, REDUCE.

Data stores

Inevitably, with such ambitious programs, the storage capacity required is very large. For practical purposes, the program has to be provided with a comprehensive knowledge of mathematics. Professor Moses estimates that MACSYMA at present uses a million 32-bit words of storage capacity but users may be able to get by with less if they make do without some of the more general functions.

Storage capacity appears to be a recurring problem in computer algebra, not merely for comprehensive systems such as MACSYMA and REDUCE, but also for more straightforward algebraic programs. The reason is easily understood. If the objective, for example, is to compute the polynomial expression $(a + bx)^n$, it is clear that the result must be a polynomial of degree n with $n + 1$ distinct coefficients. But the number of terms that arise by multiplying together one of the two elements in each of the n factors is 2^n . If n were anything like as large as, say, 100, it would be physically impossible to store the terms separately in the computer. In practice, of course, this particular computation is straightforward. But the problem of storage does compel the designers of algebraic programs constantly to be on the lookout for algorithms that make economies. For the simple binomial expansion above, for example, it plainly makes sense that the factors should be multiplied together consecutively. Nevertheless, program designers have found it necessary to provide overspill storage in case a problem should have need of it, often by means of buffer storage on magnetic tape which has then to be searched regularly and reincorporated in the accessible memory of the computer if the program so decides.

Users of both comprehensive programs such as MACSYMA and more specialized

Plenty of programs

Altogether there are probably more than 50 computer algebra systems now available. The following are the best known or most widely used.

- **MACSYMA.** Developed since 1968 at the Mathlab at the Massachusetts Institute of Technology by Professor Joel Moses and his colleagues. The objective has been a comprehensive system, capable of solving all algebraic problems. Development continues. Hitherto the system has been accessible only at MIT but the program is soon to be made available to outside purchasers. It is slower in specialized applications than purpose-designed computer programs.

- **REDUCE.** Largely developed by Professor A.C. Hearn, originally at Stanford University and then at the University of Utah. (Hearn is now at the Rand Corporation, Santa Monica.) The objective is a comprehensive system. Development has been dependent on collaboration and, collaborators say, has profited therefrom. The program is usable on a variety of machines (although some would-be users have found that adaptation is difficult). The system is still being defined and made more comprehensive. Users say that newly available programs are much more sophisticated than earlier versions.

- **SCHOONSCHIP.** Developed by Veltman at CERN in 1972 with the objective of solving problems in quantum

electrodynamics. Within its range of applications the system is fast and its demands on computer storage modest. The program is easily usable on large CDC computers but less easily on machines of smaller word length.

- **CAMAL.** The system developed at the University of Cambridge in the 1960s for solving problems in celestial mechanics. Fast for what it does and modest in demands on storage. Still widely used but now losing ground to more sophisticated systems, for example that developed by the Boeing Space Sciences Division.

- **MUMATH.** Developed by Stoutemyer at the University of Hawaii and intended for use with microprocessors. Suggested uses include teaching but also an application to small research problems, and in particular for system design.

Most computer manufacturers have algebraic programs as do the Bell Telephone Laboratories.

Such coordination as there is of academic interest in algebraic computing is undertaken in the United States by SIGSAM, a special interest group of the Association for Computing Machinery, which organizes meetings (the next at Snowbird, Utah in August this year) and which also publishes a quarterly bulletin. In Europe, coordination is through the Symbolic Mathematics Committee, organized by J.A. van Hulzen (T. Hogeschool Twente, Enschede, The Netherlands). □

systems such as SCHOONSCHIP testify to the complementarity of the two kinds of systems. The comprehensive systems are inevitably slower than the specialized systems and also less convenient to use in practice. On the other hand, access to the comprehensive systems can be of great value in the amendment of an existing program. Users also say that there are pitfalls to be avoided when programs such as REDUCE are purchased on magnetic tape and used with computing machinery different from that on which they were originally developed, which is part of MIT's justification for not hitherto making MACSYMA widely available.

Curiously enough, although the algebraic computer programs are in general demanding of computer storage, they are relatively undemanding of programming time. MACSYMA, largely the product of people working at MIT since 1968, is estimated to have taken between 50 and 100 man-years of programming effort. Other systems designed for specialized purposes can be relatively simple to design and write. One practitioner, for example, says that the time required to write a program for simple polynomial algebra is not much more than a quarter of an hour. No doubt the relatively undemanding nature of some of the programming required reflects the character of the problems to which

computer algebra has so far been applied — their intrinsic tedium. And, of course, nobody pretends that writing programs to carry out integration, to solve differential equations or to tackle problems in mathematics such as the enumeration of finite groups can be solved in a straightforward way.

What the future holds for the applications of computer algebra remains to be seen. Already, however, some of the practitioners hold that the availability of these techniques provides a means by which shortages of mathematics teachers, even at the school level, might be remedied. Indeed, the University of Hawaii has developed a computer algebra program (commercially available) that can be run on home computers, while there are ambitions to use MACSYMA in the education of undergraduates.

What educationists will have to say about this prospect is unclear. The hand calculator, once considered a threat by elementary school teachers, seems now to have been recognized as a stimulator of interest. It is harder to see how more formal computer programs will be used in teaching algebra, and it is widely thought that too indiscriminate use of computers for algebraic computations in research may blind users to the physical significance of what they are doing. □

REVIEW ARTICLE

Can dietary beta-carotene materially reduce human cancer rates?

R. Peto*, R. Doll†, J. D. Buckley‡ & M. B. Sporn§

* Imperial Cancer Research Fund Cancer Studies Unit, Nuffield Department of Clinical Medicine, Radcliffe Infirmary, Oxford OX2 6HE, UK

† Imperial Cancer Research Fund Cancer Epidemiology Unit, 9 Keble Road, Oxford OX1 3QG, UK

‡ Department of Biomathematics, University of Oxford, Oxford OX1 2JZ, UK

§ Laboratory of Chemoprevention, National Cancer Institute, Bethesda, Maryland 20205, USA

Human cancer risks are inversely correlated with (a) blood retinol and (b) dietary β -carotene. Although retinol in the blood might well be truly protective, this would be of little immediate value without discovery of the important external determinants of blood retinol which (in developed countries) do not include dietary retinol or β -carotene. If dietary β -carotene is truly protective—which could be tested by controlled trials—there are a number of theoretical mechanisms whereby it might act, some of which do not directly involve its 'provitamin A' activity.

THE possibility of discovering anticancer substances that can be prescribed rather than carcinogens that must be proscribed is attractive, for people may be more willing to accept prescription than proscription. If inhibitors of the late 'stages' of neoplastic transformation could be found, they would have the added advantage of conferring their benefits without an impracticably long latent period. Since there is growing evidence that retinoids and perhaps also carotenoids can be anti-carcinogenic, we review the intake and metabolism of 'preformed vitamin A' (retinal and retinol and its esters) and of the carotenoids (some of which exhibit 'provitamin A' activity), and we review the epidemiological evidence correlating blood levels or dietary intake of one or other of these with low cancer risks. Two independent questions will be considered.

First, within the range of variation of total blood retinol levels consistent with the avoidance of hyper- and hypo-vitaminosis A, is there any material variation in human cancer risks attributable to a protective effect either of unbound retinol or of carrier-bound retinol in the blood, and, if so, what are the important extrinsic determinants of the concentration of carrier-bound retinol in human blood? (Paradoxically, within the wide range of dietary intake rates consistent with avoidance of both hyper- and hypo-vitaminosis A, the amount of retinol derived directly or indirectly from the diet makes no material difference to the blood levels of carrier-bound retinol.)

Second, within the range of blood levels of β -carotene (or some other carotenoids) which are insufficient to cause cosmetically unacceptable skin yellowing, is there any material variation in human cancer risks attributable to a protective effect of circulating carotenoids, and, if so, what are the important extrinsic determinants of the concentration of β -carotene and other carotenoids in human blood? (In contrast with blood retinol, blood carotenoid levels are directly determined by dietary carotenoid levels, although they may also be affected by certain chronic disease states.)

The evidence relating to the above two questions will be reviewed under three headings:

(1) The qualitative difference between the pharmacodynamics of β -carotene and of retinol¹⁻³ in the human body: which

suggests that the possible relevance of these two components of 'dietary vitamin A' to cancer prevention should be assessed separately.

(2) The inhibition by retinol and various synthetic retinoids of the later, rather than the earlier, processes of carcinogenesis in experimental studies with animals and cultured cells: which suggests that if such agents have any substantial anticancer effects on humans, a reduction in risk should be measurable within about 5 years, instead of after a latent period of some decades, as might be expected for inhibitors of the 'early' processes of carcinogenesis.

(3) The tendency, observed in many epidemiological studies, for some human cancer risks to be inversely associated with blood retinol, and with blood or dietary β -carotene.

Pharmacodynamics

'Vitamin A' is a nonspecific term embracing two different families of dietary factors, one family comprising the various types of 'preformed vitamin A' (chiefly retinyl esters, but also retinol and retinal), and the other family comprising the various types of 'provitamin A' (chiefly β -carotene, but also those other carotenoids that can be metabolic precursors of retinol). If retinal (the aldehyde $C_{20}H_{28}O$) is written $X=O$, then retinol, the corresponding alcohol, is $H-X-OH$, and β -carotene is the polyunsaturated hydrocarbon $X=X$, which may be oxidized by the enzyme dioxygenase to two molecules of retinal, chiefly in the intestine but perhaps also in the liver⁴.

The various types of preformed vitamin A are converted efficiently by the human body into retinol, and a variable fraction of dietary β -carotene is converted in the human intestine into retinal. Nutritionists estimate² that even among poorly nourished people β -carotene usually yields only about one-sixth of its weight of retinol; furthermore, the fraction converted decreases as the amount of β -carotene in the diet increases so that even massive doses of β -carotene will not yield enough intestine-derived retinol to cause systemic toxicity ('hypervitaminosis A'), unless massive amounts of preformed vitamin A are also taken. Despite this inefficient intestinal conversion of β -carotene, and the even less efficient intestinal

Caution

Unwary readers (if such there are) should not take the accompanying article as a sign that the consumption of large quantities of carrots (or other major dietary sources of β -carotene) is necessarily protective against cancer, and the correlation between blood retinol and cancer avoidance is, for the time being, *sub judice*.—Editor, *Nature*.

Table 1 Prospective studies of cancer in relation to total blood retinol in stored blood samples taken from apparently healthy people

Country	Relative risks (in sequence from low to high blood retinol quintile)					No. of cancers (all sites)
	5.3:	2.9:	3.8:	3.0:	1*	
Georgia, USA ³¹	5.3:	2.9:	3.8:	3.0:	1*	85
United Kingdom ³²	2.2:	1.7:	1.6:	1.1:	1	86

* Top quintile based on only seven cancers, and so unreliable.

conversion of most other carotenoids, many people in rich countries and most people in poor countries derive the majority of their retinol from the provitamin A carotenoids in plants. However, some of the dietary β -carotene and other carotenoids is absorbed unchanged from the intestine, circulates around the body, and, being fat-soluble, is stored in adipose tissue. Various factors then affect its metabolism and excretion (for example, diabetes⁵, thyroid disease⁵ and, paradoxically, anorexia nervosa^{6,7} will elevate blood β -carotene levels), but other things being equal blood and adipose tissue levels of β -carotene and other carotenoids are directly related to the amounts ingested over the preceding few weeks⁸. Normally, a daily average of only a few milligrammes of β -carotene is consumed. This makes human fat appear yellow but is not sufficient to cause noticeable yellowing of the skin. However, if a few dozen milligrammes of β -carotene per day are consumed for several weeks by Caucasians, 'hypercarotenaemia' may develop, that is, skin yellowing may become cosmetically noticeable. No medically significant side effects of hypercarotenaemia are known⁹, and the condition disappears within a few weeks of reversion to a normal diet.

The mechanisms for storing, transporting and excreting retinol are more specific than those thus far discovered for β -carotene. Intestine-derived retinol enters the blood as 'unbound' retinol, esterified but not bound to the specific carrier protein for retinol (see below). Unbound retinol is rapidly absorbed from the blood by the liver, where it is stored. To transport controlled amounts of retinol from the liver to the peripheral tissues where it is needed, the liver synthesizes a carrier protein called retinol binding protein (RBP) at a rate which, unless the liver's stores of retinol are almost completely depleted, is not strongly dependent on the amount of retinol currently stored in the liver. (The total rate of loss of retinol from the liver is, however, directly dependent on the total amount stored in the liver^{10,11}, so other, presumably degradative, pathways must also exist, and will provide some homeostatic control against unlimited accumulation of retinol in the liver.)

RBP molecules can leave the liver as soon as they have picked up a retinol molecule from the liver stores; but if the liver stores are severely depleted RBP builds up in the liver, waiting for some unbound retinol to be derived from the intestine. This RBP build-up is characteristic of hypovitaminosis A, that is, vitamin A deficiency. At the opposite extreme, if due to regular heavy intake of preformed vitamin A the rate of derivation of unbound retinol from the intestine consistently exceeds the maximum possible rate at which the liver can get rid of retinol (both by steady carrier-bound release and by concentration-dependent degradation), the liver's retinol storage capacity will eventually be exceeded, damaging the liver and allowing significant amounts of unbound retinol (not bound to RBP) to circulate. This may lead to serious symptoms of hypervitaminosis A, since unbound retinol should never circulate in significant quantities except in brief transit from the intestine to the liver. However, most of the population of developed countries has a diet-derived intake of retinol within the wide range which avoids the well-defined physiological conditions of hypovitaminosis A (unbound RBP in the liver) and hypervitaminosis A (unbound retinol in the blood). Within this wide range, blood levels of carrier-bound retinol (retinol bound to RBP) are not materially dependent on the amount of free retinol derived from the intestine. For example, among Swedish acne patients an

increase of the normal dietary dose of retinol by a few thousand per cent for a few weeks increased carrier-bound retinol by only about 10% (ref. 12). Likewise, in New York, liver retinol levels, which are directly related to dietary intake rates, are uncorrelated with blood retinol levels¹³.

However, blood levels of carrier-bound retinol do have extrinsic determinants. For example, oral contraceptive users have substantially elevated blood retinol levels^{14,15}, and there are presumably also various other hormonal or non-hormonal determinants of the consistent differences between people in their long-term average carrier-bound retinol levels (including¹⁶ diseases of the liver, thyroid and kidney), although none seem to be well understood. The key finding for this whole review is that for most people in developed countries the amount of intestine-derived retinol is not one of the important determinants of the blood levels of carrier-bound retinol (or of total retinol, since free retinol is usually present in the blood only at much lower concentrations than carrier-bound retinol is).

Possible short latency of antitumour effects of retinoids and related compounds: The ICRDB¹⁷ have compiled summaries of some 324 cancer research papers concerned with retinol, other retinoids, β -carotene or other carotenoids, and the October 1979 issue of *Federation Proceedings* contains several further relevant papers. Together, these show that various retinoids can reversibly suppress the malignant behaviour of cultured cells that have been transformed by viruses, chemicals or ionizing radiation; they can delay or prevent the onset of cancer in experimental animals that have been treated previously with DNA-binding carcinogens; they can cause human skin keratoses or oral leukoplakias to regress; they can delay the appearance of transplanted tumours; they can prevent 'tumour promoters' from eliciting papillomas on mouse skin; and they can inhibit some of the very specific biochemical effects of tumour promoters on cells *in vivo* or in culture. Although inappropriate use of retinoids may do more harm than good¹⁸, it thus seems reasonable to hope that exposure of human tissues to retinoid-like activity can be manipulated to reduce cancer risks¹⁹.

Neoplastic transformation can often be subdivided into 'stages', the early stage(s) involving only partial alteration of a normal cell and the later stage(s) involving conversion of a partially altered cell into a fully neoplastic cell. Finally, to be effective the neoplastic cell must not be eliminated, and must proliferate into a pathological tumour. Early- and late-stage changes are both necessary (so halving either might be an equally effective way of halving cancer risks), but typically have different causes^{20,21}. The demonstrated protective effects of retinol (and, perhaps, β -carotene) are such as to suggest that it is the later stages of neoplastic progression which are chiefly affected, together perhaps with the final process of tumour growth. There might often be a latent period of a few decades before a doubling or halving of the early processes had any material effect on human cancer risks, but, to judge by the rapidity with which the effects of stopping smoking are evident, a doubling or halving of the late processes might have noticeable effects on human cancer risks within about 5 years, while effects on the final process (proliferation) might be even more rapidly evident. So, human cancer risks might be noticeably influenced by the average levels of blood retinol within a period of as little as 5 or 10 years. This conclusion greatly facilitates observational epidemiology and may even make practicable a randomized human evaluation of the anticancer effects of such agents (or of any other agents which inhibit either the late or the final processes).

By contrast with the retinoids, the question of whether dietary β -carotene or some other carotenoids can affect cancer risks has not received nearly as much experimental attention. Possible mechanisms to consider include (1) a direct retinoid-like effect of some carotenoid(s) on cellular differentiation in the target tissues, (2) conversion in the target tissue of some carotenoid(s) into molecules with some such retinoid-like effects, or (3) protection by carotenoids of the target tissues via mechanisms not

related to control of cellular differentiation (for example, by enhancing some immunological function^{22,23}, or by quenching singlet oxygen, a highly reactive excited molecular species which may be generated as a toxic by-product of many normal metabolic processes in both animals and plants: indeed, β -carotene is the most efficient quencher of singlet oxygen thus far discovered²⁴, and in plants the chief function of carotenoids may be to quench the singlet oxygen produced by photosynthesis²⁵). Before possible mechanisms are discussed, however, it must be asked whether any carotenoids have any general anticancer effects whatever.

We know of only nine experiments on the modifying effects of systemic β -carotene²⁶⁻²⁸ or of carrot extracts²⁹ on induced tumours. In each experiment a protective effect was claimed, but lack of published details of the exact experimental circumstances makes some of these claims difficult to evaluate. If the reported²⁸ protective effects of dietary β -carotene against tumour induction by transplanted cells or viruses are confirmed, this may merely reflect the reported^{22,23} enhancement of certain immune parameters by dietary β -carotene, the relevance of which to human cancer remains uncertain.

If the reported protective effects of systemic β -carotene against chemically induced skin cancers^{27,29} are confirmed (and are not trivially attributable to physical changes in the skin) this would be of more interest than protection by systemic β -carotene against skin cancer induction by UV light^{26,27}, for the latter may at least in part reflect a photoprotective rather than an anticancer effect. Further studies in all these areas will be needed before any general protective effects of β -carotene in experimental animals can be accepted. (No excess of tumours was reported in the chronic toxicity test of β -carotene³⁰, but details were not provided.)

Epidemiology

The foregoing laboratory evidence has been introduced not because it proves that β -carotene will have an anticancer effect, but chiefly as background material to aid interpretation of the epidemiological observation that people with above average blood retinol levels or above average β -carotene intake rates are at below average risk of cancer. Although the available epidemiological data (see below) indicate that both factors are associated with low cancer risk, this does not constitute overwhelming evidence for any real protective effect, for the differences are not great. A moderate relative risk (for example, a twofold difference in cancer risk between the bottom third and the top third of some population with respect to carrier-bound retinol in the blood or β -carotene in the diet) might reflect a

truly protective effect, but it might arise indirectly merely because carrier-bound retinol levels are correlated with some other factor that is a real determinant of cancer risk. Even more plausibly, perhaps people who like to eat a lot of vegetables get both β -carotene and some truly protective factor from them, or tend to avoid some real cause of cancer in meat or other foodstuffs of which they consume less. These difficulties of interpretation can be reduced by seeking epidemiological evidence for a protective effect in many different studies in different types of population but, even if there is a medically significant but numerically moderate protective effect of dietary β -carotene, and even if these epidemiological studies nearly all estimate it reasonably accurately, we may still not be sure that they have done so unless the effect can be reproduced experimentally in randomly controlled prospective intervention studies among humans.

Blood retinol and cancer avoidance

The differences between total retinol and carrier-bound retinol concentrations in human blood are often smaller than the random error of measurement in blood retinol estimation, and so we shall not distinguish carefully between them in some of what follows.

In general, because of the wide range of blood levels of retinol in countries where vitamin A deficiency is common, any true protective effect may show up most strikingly in malnourished populations. However, the relevance of carrier-bound retinol and of β -carotene can be assessed independently only in well-nourished populations where variation in β -carotene intake has little effect on carrier-bound retinol levels and so the blood levels of the two factors are largely uncorrelated.

Prospective serum studies: Retinol can be reasonably stable in stored blood samples, and two studies^{31,32} have analysed the total retinol levels in stored samples of blood taken some years previously for unrelated research purposes from American and British people who were apparently free of cancer at that time. In both studies, people with total blood retinol levels near the top of the normal range were at significantly lower risk of developing cancer over the subsequent few years than those with levels near the bottom of the normal range (Table 1). But, although the apparent protective effects are statistically significant, so few cancers were observed that there is substantial statistical uncertainty in the magnitude of the effect. (Anyone who has been sitting for a few years on a few thousand blood samples might consider matching them up with the local cancer registry or with official death statistics, as long as it is still possible to measure the retinol and, ideally, the RBP and β -carotene in them.)

Against this, one preliminary observation which, if confirmed, suggests that retinol is not importantly protective relates to oral contraceptive (OC) use. OC users have substantially elevated levels of blood retinol^{14,15}, but the aggregated evidence from three studies of OC use in relation to the onset of various diseases, including all cancers, does not suggest that any material reduction in cancer risk is associated with 5-9 or even 10-14 years of continuous use of OCs. Tumours of the breast, endocrine, reproductive and genital organs are excluded (as these are likely to be directly dependent on hormonal factors or on sexual activity) and smoking habits are standardized for (since OC users tend to smoke more than non-users). A non-significant deficit of remaining tumour onsets among OC users in the Oxford (M. P. Vessey, personal communication and ref. 82) and Royal College of General Practitioners (C. R. Kay, personal communication and ref. 83) studies of a total of 60,000 women is counterbalanced by an apparent excess of skin cancer onsets among OC users in the American Nurses Health Study of 120,000 women (F. E. Speizer, personal communication; unfortunately, these American data could not be corrected for any hypothetical biases due to a correlation of sunbathing with OC use). Lack of any protective effect among long-term OC users would be quite strong evidence that blood retinol is not importantly protective, unless the rise in blood retinol among

Table 2 Case/control studies of blood retinol or β -carotene

Country	Mean values (mg l ⁻¹)		Substance	No. and site(s) of cancers
	cases	controls		
India ³³	0.31	0.48	Retinol	555 oral
India ³³	0.68	0.85	β -carotene	555 oral
Pakistan ³⁴	0.22	0.40	Retinol	203 oral
Pakistan ³⁴	0.40	0.62	β -carotene	203 oral
East Africa ³⁵	—*	—*	β -carotene	17 naso-pharyngeal
United Kingdom ³⁶	0.55	0.59	Retinol	35, many sites
United Kingdom ³⁶	0.83	0.92	β -carotene	35, many sites
United Kingdom ³⁷	0.46	0.66	Retinol	28 lung
United Kingdom ³⁸	0.47	0.60	Retinol	26 lung
United Kingdom ³⁸	1.05	1.26	β -carotene	26 lung
United Kingdom ³⁹	0.84	1.34	'Vitamin A'†	245, many sites
United States ⁴⁰	0.25	0.48	Retinol	51 gastro-intestinal
United States ⁴⁰	1.4	2.0	β -carotene	51 gastro-intestinal

* 'Carotene significantly lower' (in cases than in controls), but quantitative information not provided.

† Analytical methods not specified.

OC users is somehow due to a decrease in the ease with which peripheral cells can utilize retinol (or is in some other way compensatory).

Further large prospective studies of blood retinol are still needed. If they confirm the inverse association with cancer risks then identification of various intrinsic or extrinsic determinants of blood retinol will be necessary before it can be decided whether retinol *per se* is protective. Ideally, one might manipulate the determinants of blood retinol in a large randomized trial of cancer prevention, but as this may be impracticable (depending on the nature of these determinants) even uncontrolled observation, as in the above three studies of OC users, may provide useful clues.

Retrospective serum studies: Comparison of the total blood retinol levels of recently diagnosed cancer patients with the total blood retinol levels of control patients is easier and quicker than a prospective cohort study. Of the nine studies in India³³, Pakistan³⁴, East Africa³⁵, Britain³⁶⁻³⁹, or the United States^{40,41}, the eight which had at least a small control group³³⁻⁴⁰ all reported lower total retinol levels in cases than in controls (Table 2: all but one³⁶ of these eight are significant).

It may be relevant to the succeeding section to note that blood levels of β -carotene were also measured in some of these case/control studies. Among Indian, Pakistani and East African cases blood β -carotene levels were significantly lower than in controls. Unfortunately, β -carotene was measured in only three of the Western studies³⁸⁻⁴⁰, in each of which it was also somewhat lower in cases than in controls, even though no particular correlation between blood retinol and blood β -carotene levels need exist in developed countries.

The evidence provided by these eight case/control studies for either inverse association existing is not reliable, however, for in some studies there is no guarantee that the controls were appropriate and in all the studies the results might have been biased by a (hypothetical) 'nuisance effect' of depression of β -carotene or carrier-bound retinol by large, established cancers. There is still a need for large case/control studies in both developed and malnourished countries of carrier-bound retinol (and β -carotene) levels in people with small tumours that are unlikely to have had any material effect on diet or general metabolism (or in people who have been cured of malignant tumours some time previously with no obvious permanent damage to their digestive system or metabolism).

Dietary β -carotene and cancer avoidance

Turning from the blood to the diet, and shifting our emphasis from retinoids to carotenoids, a variety of independent epidemiological investigations have, in recent years, reported that cancer risks are inversely related to consumption of 'vitamin A'. This term embraces both retinol and the provitamin A carotenoids; but the total amount of 'vitamin A' in human diets from β -carotene (plus other carotenoids) often exceeds that from retinol.

In many cultures one particular plant or group of plants accounts for the majority of the dietary β -carotene: for example, carrots in North America⁴², yellow-green vegetables in Japan⁴³, red palm oil in West Africa⁴⁴ and dark green leafy vegetables among Singapore Chinese⁴⁵. Many epidemiological studies have investigated the relationship between β -carotene and cancer almost by accident, because dietary questionnaires designed for more general purposes happened to ask about the main local β -carotene source and it was realized only later that the lower cancer risks found among users of this particular foodstuff might be due to its β -carotene content. The correlation between the average β -carotene intake rates of different individuals over the past few years and their intake as assessed by such questionnaires cannot be high. So, if the observed 1.5- or two-fold relative risks (RR) of cancer typically observed when comparing low and high β -carotene intake groups (as assessed by dietary questionnaires) are indeed due to the β -carotene intake, the relative risk comparing true low with true high intake must presumably be somewhat more than

twofold, and might in principle be substantially greater. The results of these studies (5 prospective and 15 retrospective) are summarized in Table 3, and discussed below.

Prospective dietary questionnaires: All the Japanese in certain administrative areas were asked⁴³ to complete an extensive dietary questionnaire in 1965, and 265,118 did so. Over the next decade 807 died of lung cancer, yielding an age/sex/smoking-standardized relative risk of 1.4 ($P < 0.01$) for those who ate vegetables containing more than 1,000 IU of β -carotene per 100 g 'less than daily' with those who ate them 'daily'. Significant negative associations with daily use of these vegetables were also reported for stomach and prostate cancer, but the findings for other cancers were not reported. These yellow-green vegetables account for more than half of the β -carotene in the Japanese diet.

8,278 Norwegian men who had replied to postal questionnaires about both smoking (in 1964) and diet (in 1967) were followed for 5 years⁴⁶. A 'vitamin A index' (chiefly, in fact, a β -carotene index, since it included carrots and green vegetables but not liver or vitamin pills) was devised, low values of which were predictive of lung cancer incidence ($RR = 3$ for people in the bottom third of the population with respect to the 'vitamin A index' compared with people of similar age, sex and smoking habits in the top two-thirds; $P = 0.01$, but based on only 36 cases). A total of some 40,000 people completed the same questionnaire in 1966-67 in Norway or Minnesota⁴⁷, and the 10-year follow-up of these is being analysed, with apparently promising results (E. Bjelke, personal communication).

No other large prospective studies of either dietary β -carotene or preformed vitamin A have been published. It would be helpful if anyone who can link dietary questionnaires administered for any reason some years ago to subsequent cancer mortality (or, preferably, incidence) would do so, if they have reasonable estimates of consumption of at least the one or two main local β -carotene-containing plants plus at least 100,000 person-years of follow-up. For example, in Finland the detailed dietary questionnaires administered by the Research Institute for Social Security during 1966-72 to a geographically defined sample of 20,000 adults could be linked to the few hundred subsequent cancer onsets. Likewise, the Japan-Hawaii cancer project has by now sufficient data to provide some useful prospective information, although in Hawaii, as in Japan⁴³, one must interpret with particular caution correlations between any dietary habit and any cancer if both are subject to substantial trends, for artefacts will arise if the more modern members of society are more affected by both trends than are the more conservative members. Finally, dietary questionnaires have been administered to various large populations for research into cardiovascular disease, or for nutritional research, and these too might be exploitable. (Preliminary analyses of the 19-year follow-up of 2,000 people in Chicago who completed a 28-day dietary diary in 1959 indicate an inverse association of total cancer risk with β -carotene intake but not with retinal intake: R. Shekelle, personal communication.)

Retrospective dietary questionnaires: These involve asking people with cancer ('cases') and people without cancer ('controls') what vegetables, vitamin pills, etc. they normally ate in previous years, from which can be estimated⁴⁸ the relative risk of cancer in people with low, as opposed to high, β -carotene intake (or retinol intake—a nonsignificantly lower risk has been reported among users of vitamin A pills^{49,50}). One general difficulty is that feeling ill or knowing they have cancer may cause people to describe their past diet somewhat differently from how they would have described it if they had been healthy controls. If this bias is not substantial, the enormous advantage of the retrospective method is that it is quick (you do not have to wait for 10 years to see who develops the disease) and much less work (one questionnaire per case plus one per control), which allows more patients with cancer to be studied with consequently less random error in the final relative risk estimation. Although simple statistical methods⁴⁸ exist for correcting for age, sex, smoking and any dietary confounding variables

before calculating a test for a trend in the relative risk of cancer with respect to β -carotene or retinol intake, unfortunately investigators have often failed to use them or have used them but have failed to test for a *trend*, on only 1 degree of freedom.

At Roswell Park Memorial Institute in the United States, a few thousand cancer patients and a few thousand controls completed dietary questionnaires during 1957–62. Milk and some but not all other dietary sources of retinol were assessed, and carrots and some other major sources of β -carotene were assessed. A 'vitamin A' estimate was constructed for each person, which chiefly reflected provitamin A (β -carotene) intake, since on average five-eighths came from carrots (as β -carotene) and some more came from other vegetable sources of β -carotene⁴². Subjects were then classified as high, middle or low with respect to this estimate. When the 292 lung cancer patients were compared with selected controls⁵¹, the age/sex/smoking-standardized cancer risks relative to the high intake group were 1.0, 1.5 (middle) or 1.7 (low), and a test of the published data for an age/sex/smoking-standardized trend yields $P < 0.001$. An IARC case/control study of 233 Singapore Chinese lung cancers⁴⁵ indicated a $P < 0.01$ twofold protective effect of eating twice or more per week 'dark green leafy vegetables' (for example, kale), which are a major source of β -carotene for these people, but a similar case/control study⁵⁰ of 104 British lung cancers indicated no protective effect ($RR = 1.0$) of dietary 'vitamin A'. Returning to Roswell Park, the 569 bladder⁴², 122 oesophagus⁵² and 421 larynx⁵³ cancer patients were compared with selected controls. In seven sub-groups of decreasing (largely pro-) vitamin A intake, the risks of bladder cancer relative to those of the highest intake group were 1.0, 1.2, 1.4, 1.2, 1.7, 1.9 and 2.1, and a test of these published data for a trend again yields $P < 0.001$. Among the 122 oesophageal cancer patients divided into three sub-groups by dietary β -carotene intake relative risks of 1.0, 1.6 and 1.9 were found ($P = 0.03$ for trend), while among the 421 male larynx cancer patients divided into four dietary sub-groups, relative risks of 1.0, 2.1, 1.9 and 3.0 were found ($P < 0.005$ for trend) after adjustment for age, tobacco and alcohol. (Similar adjustment of the data for 581 oral cancers is necessary before the twofold protective effect seen in an unadjusted analysis⁵⁴ can be interpreted.)

The Roswell Park questionnaires comprise the largest set of case/control data available anywhere, and it is to be hoped that a single analysis of all cancers separately and together in relation to all controls will eventually become available in which trends in relative risks with respect to fine subdivisions of β -carotene and perhaps of retinol intake will be tabulated and standardized for finer subdivisions of age and smoking than heretofore.

There are also several studies relating gastrointestinal (GI) cancer risks to low vegetable intake rates but in all of these it is possible that vegetable components other than β -carotene (or dietary components other than the vegetables) account for the apparent protective effects in the GI tract.

Case/control studies of GI cancer have been reported from Hawaii^{55,56} and Israel^{57,58}, but in neither case were analyses reported with respect to specific estimates of β -carotene or retinol intake. Recently, the Israeli data on 406 cases and 812 controls have been re-analysed⁵⁹ with respect to an estimate of total dietary β -carotene for each case and for each control based on separate replies about more than 200 foodstuffs. There was no significant correlation between the risk of GI cancer and the estimated daily dietary carotene intake here (or in preliminary analyses of the Roswell Park data on GI cancers: S. Graham, personal communication). The only other study of GI cancer which has sought but not found any apparent protective effect whatever is the IARC case/control study of oesophageal cancer in French males⁶⁰, where the mean β -carotene (and retinol) intake was the same for the 200 cases as for the controls. By contrast, the IARC studies on oesophageal cancer in Iran⁶¹ indicate that the high-incidence areas were characterized by a diet extremely poor in green vegetables, and the main dietary finding in the subsequent IARC case/control studies within

high-incidence villages of several dozen dietary variables was a systematic lack of fresh fruit and vegetables among the cases⁶², although specific estimates of β -carotene (and retinol) intake in cases and controls were not published.

At other GI sites four separate case/control studies have been reported⁴⁷ each of which shows an age/sex-standardized relative risk of cancer that was statistically significantly greater than unity in people whose 'vegetable index' was lower, as opposed to higher, than normal. (Note, however, that only two control series, one in Norway and one in the United States, were interviewed.) The findings in these studies were⁴⁷:

- | | | |
|--------------------|-------------------------|------------|
| (a) Norway, | 228 stomach cancers: | $RR = 1.7$ |
| (b) United States, | 83 stomach cancers: | $RR = 1.5$ |
| (c) Norway, | 278 colorectal cancers: | $RR = 1.4$ |
| (d) United States, | 373 colorectal cancers: | $RR = 1.4$ |

In Liverpool, UK⁶³, a 1.6-fold relative risk of GI cancer ($P < 0.01$) was found if green vegetables were eaten less than weekly. Among vegetarians⁶⁴, a nonsignificant twofold relative risk was likewise found for colorectal cancer in Seventh Day Adventists who eat green, leafy vegetables 'less than' compared with 'at least' weekly. In India, it is reported⁶⁵ that studies of vegetarian sects consistently reveal smaller GI cancer risks, but no supporting data or references are provided. (Studies relating β -carotene intake to cancer risks within vegetarian sects would be free of confounding by associations with meat consumption.) Finally, in case/control studies⁶⁶ in a total of 521 gastric cancer patients in four countries, a lower intake of fresh fruits and vegetables in cases than in controls is indicated in the discussion, but the control data are not given (nor is the percentage of the controls who themselves had some other form of cancer).

Future research directions

Laboratory research: Most epidemiologists are well aware that the human evidence thus far available (Tables 2, 3) falls short of proof that increasing the average β -carotene intake would be of any material benefit in rich countries (or even in poor ones). *A fortiori*, the foregoing epidemiological studies may fail to interest many laboratory research workers whose habit is to trust only striking differences in well-controlled experiments. However, the perfect may well be the enemy of the relevant, for the detailed understanding of mechanisms which is slowly emerging from laboratory investigations of a few particular chemical, viral or physical agents (for example, SV40, phorbol esters or 2-acetylaminofluorene) may be of little immediate value if few human cancers are in reality due to these particular agents. Although in the long term fundamental research into any area of cell biology may prove relevant to the prevention or treatment of cancer, progress towards important preventive measures may be more rapid if, in addition to fundamental research, laboratory workers take the inevitably inconclusive leads provided by epidemiologists reasonably seriously (and vice versa, of course), one ultimate aim of this collaboration being to specify a few promising hypotheses in sufficient detail to make practicable their randomized evaluation among several thousand apparently healthy people for 5 or 10 years. Several laboratory studies are now possible that would test the plausibility of the more obvious mechanisms whereby β -carotene might be beneficial.

First, and most obvious, is to repeat the animal experiments in which important anticancer effects of β -carotene have been reported^{26–29}. Mechanistically, it would be interesting to discover whether dietary (not topical) β -carotene, started some weeks after weak (not massive) initiation of animals with DNA-binding carcinogens, has any material inhibitory effects on the later stage(s) of carcinogenesis, especially in tissues where dietary β -carotene is known to accumulate⁶⁷ or to cause accumulation of retinol⁶⁷.

Second, does β -carotene itself have any of the many direct effects on cultured cells or tissues that retinoids have? Such effects are relatively quick to obtain and easy to seek, although a variety of cell culture conditions and β -carotene vehicles may

Table 3 Questionnaire studies of cancer in relation to intake of some β -carotene-rich vegetable(s), or of 'vitamin A'. Note that 'Vitamin A' was estimated from a limited range of dietary components, chief among which in most of these studies has been the β -carotene in certain vegetables

Country	Relative risks (lower: higher intake)	No. of cancers	Subdivision by
Norway ⁴⁶	2.6: 1	36 lung	'Vitamin A'
Norway ⁴⁷	1.7: 1	228 stomach	Vegetables
Norway ⁴⁷	1.4: 1	278 colorectal	Vegetables
Minnesota, USA ⁴⁷	1.5: 1	83 stomach	Vegetables
Minnesota, USA ⁴⁷	1.4: 1	373 colorectal	Vegetables
Buffalo, USA ⁵¹	1.7: 1.5: 1	292 lung	'Vitamin A'
Buffalo, USA ⁴²	2.1: 1.9: 1.7: 1.2: 1.4: 1.2: 1	569 bladder	'Vitamin A'
Buffalo, USA ⁵²	1.9: 1.6: 1	122 oesophagus	'Vitamin A'
Buffalo, USA ⁵³	3.0: 1.9: 2.1: 1	421 larynx	'Vitamin A'
Singapore ⁴⁵	2.2: 1	233 lung	Green vegetables
United Kingdom ⁵⁰	0.7*: 1.3: 1.4: 1.1: 1	104 lung	'Vitamin A'
Israel ⁵⁹	1.1: 1.1: 1	406 gastrointestinal	β -carotene†
France ⁶⁰	1.0: 1	200 oesophagus	β -carotene
United Kingdom ⁶³	1.7: 1.1: 1	514 gastrointestinal	Green vegetables
US vegetarians ⁶⁴	2.2: 0.9: 1	41 colon	Green vegetables
Iran ⁶²	1.7: 1	344 oesophagus	Raw vegetables
Japan ⁴³	1.3: 1	611 male lung	"green/yellow‡" vegetables
Japan ⁴³	1.5: 1	196 female lung	
Japan ⁴³	2.0: 1	63 prostate	
Japan ⁴³	1.1: 1	1823 stomach	

Note: In addition, a US prospective study in 1948–54 reported an inverse association (risk ratios 2.2:1.6:1) with all-causes mortality⁸⁴.

* Bottom quintile based on only seven cancers, and so unreliable.

† Relative risks estimated from simultaneous regression on both retinol and β -carotene intake.

‡ Defined as any which contained over 1,000 IU of carotene per 100 g.

have to be studied. The weak effects already reported⁶⁸ could be pursued.

Third, although much of the enzymic oxidation of β -carotene to retinal takes place in the intestine, is there sufficient oxidation at other sites in humans to provide an appreciable source of localized retinol, or some other substance with retinoid-like activity? Dioxygenase, the enzyme which can oxidize β -carotene to retinal, has been demonstrated chiefly in the intestine, but it is also present in the liver and, to a lesser extent, the kidneys⁴. In which human tissues can β -carotene be converted into retinol? (When mice are treated with β -carotene deposits of it accumulate in many tissues, and deposits of retinol also accumulate in the lung and skin⁶⁷.)

Fourth, β -carotene is among the most efficient substances known for quenching the excitation energy of singlet oxygen^{24,25,69} and also for trapping certain organic free radicals⁷⁰. It has been suggested that anticancer effects (albeit perhaps on only the 'early' stage(s) of carcinogenesis) may be produced by deactivation of these chemical species⁷¹ or by otherwise preventing the lipid peroxidation or other oxidative damage which they might cause^{72–75}. This suggestion could, of course, be tested if people with a deficient ability to 'quench' singlet oxygen or to 'trap' certain types of free radical could be identified and their cancer onset rates monitored. If either chemical species is relevant, can realistic physiological concentrations of β -carotene, or any other carotenoids, effect any material reduction in the amounts of lipid peroxidation or other cellular damage caused by free radicals or by unquenched singlet oxygen, and is such oxidative damage importantly involved in carcinogenesis?

Fifth, both in poor and in rich countries, what are the chief dietary and other determinants of the differences that now exist between the blood levels of carrier-bound retinol and of β -carotene in different people? To what extent, for example, do they depend on dietary β -carotene, intestine-derived retinol, other micronutrients, or hormonal status? Comparing poor countries with each other⁷⁶, there is a correlation between dietary (pro-) vitamin A and blood retinol; but, as one would expect, the relationship seems to reach a maximum and then flatten off. Both oral contraceptives and pregnancy strongly affect blood retinol levels^{14,15}; what of menarche, menopause and dietary or medicinal hormones? Finally, diseases of the kidney¹⁶, thyroid^{5,16} and pancreas⁷ have been reported to cause increases or decreases in blood retinol or β -carotene; how do

they do so, and which other diseases (especially of hormone-secreting tissues) do so?

This list of laboratory questions could doubtless be extended. The general point is that the evidence linking β -carotene intake to cancer risk in man is already sufficiently consistent for laboratory workers to start to seek plausible explanations for it, bearing in mind the superficially paradoxical fact that dietary β -carotene may not affect serum retinol very much in developed countries.

Observational epidemiology: There are some populations (notably in parts of West Africa and Brazil) within which β -carotene intake varies enormously, due to the use of red palm oil. If β -carotene is really protective this wide variation should produce quite a striking variation in the relative risk of cancer, and so the Imperial Cancer Research Fund are supporting a case/control study (with J.D.B. and Dr. A. Kalache) in a part of Brazil where red palm oil is used by some, but not by all. Apart from this, and a general refining of questionnaires and statistical analyses to pick out the separate effects of β -carotene and retinol more specifically, we need more rigorous serum epidemiology to investigate the predictive value of serum retinol and β -carotene retrospectively or, preferably, prospectively. (Repeat measurements of retinol and β -carotene a year or more apart on a number of individuals to estimate within-person and between-person variability is a prerequisite for proper interpretation of such studies.) Future serum-based case/control studies are needed both in developed countries and in populations where hypovitaminosis A is endemic, but these should be large and should exclude artefactual effects of established tumours on blood levels (via effects on recent appetite or on general metabolism), perhaps by restricting interest to pre-cancerous lesions, to small non-digestive organ tumours, or to small tumours which have been cured without much trauma a year or more previously. Better, however, would be truly prospective studies, either starting *de novo* or utilizing blood samples taken years ago for other purposes but stored in conditions that still allow estimation of retinol and, ideally, β -carotene. Also, the question of whether prolonged oral contraceptive use reduces cancer risks should be pursued, as this may help determine whether any inverse association between blood retinol and cancer is due to a true protective effect of blood retinol, or merely to a secondary association (in which case, with what?). Likewise, any other extrinsic or intrinsic determinants of blood

retinol deserve epidemiological study. Finally, to help elucidate mechanisms, cancer onset rates might be monitored among small populations (if these can be found) with chronic metabolic defects which markedly elevate or depress blood levels of unbound retinol, carrier-bound retinol or β -carotene, or among people deficient in the β -carotene-oxidizing enzyme dioxygenase, in some of the normal mechanisms for quenching singlet oxygen, or in the enzymes for recognizing or utilizing retinol in peripheral tissues.

Intervention studies: In view of the general association previously observed between low vegetable use and cancer risks, it must be assumed that future, perhaps more specific, observational studies of serum β -carotene and retinol levels are in aggregate likely to continue to yield some inverse association with cancer risks, although the magnitude of the relative risk that will emerge cannot be predicted. If it is large then a clear need for randomized intervention studies will arise, but even if it is only moderate there will be no reliable way other than by randomized intervention to decide whether it reflects a cause-and-effect relationship. However, although analogy with the effects of giving up smoking suggests that inhibition of the late stages of human neoplastic progression should yield measurable effects within about 5 years, there may be little benefit apparent during the first 2 or 3 years of such a trial. Consequently, if randomized intervention trials are eventually to be undertaken then the sooner they are started the better, as long as they test β -carotene (preferably at as high a daily dose as possible without significant skin yellowing) rather than retinol, in the hope of detecting either any benefits of intestine-derived retinol or any direct benefits from circulating or adipose tissue β -carotene.

Testing β -carotene has the advantage of avoiding the risk of hypervitaminosis A arising among a few subjects who choose to overdose themselves⁵, for 'as far as is known hypercarotenosis in all its forms is entirely harmless'⁹ (and there is no evidence that the one man reputed to have died of carrot juice actually did so⁷⁷). Indeed, the joint FAO/WHO Expert Committee on Food Additives⁷⁸ formally estimated the acceptable daily intake of β -carotene for a 70 kg adult to be 350 mg per day! Since the normal range of human intake of this biologically active substance is only a few mg per day, this committee showed a remarkable faith in their ability to extrapolate from apparently complete safety in a few rats to safe levels in large human populations, but amounts of one or two hundred mg per day are regularly prescribed for the treatment of erythropoietic protoporphyria without causing hypervitaminosis A, liver dysfunction or any other apparent ill effect^{8,79-81} and large amounts are also consumed by many West Africans and Brazilians who prepare food with red palm oil. (Ghanaian liver retinol and blood carotene levels are grossly elevated, but their median blood retinol is normal⁴⁴.)

Long-term randomized intervention studies with β -carotene might be worthwhile on almost any category of persons at high enough risk of cancer for several dozen occurrences, recurrences or progressions to be anticipated in a practicable sized study group within a few years. However, among really high-risk people trials of synthetic retinoids may be preferred, despite the possibility of toxic effects. With β -carotene, perhaps the most interesting possibility is randomized intervention in the general population, or in those of the general population who seem to have less retinol or β -carotene than average by dietary questionnaires or by serum analysis. It may be most practical in such studies to restrict interest to people over 50 or over 60 years old, because in younger people (even if they smoke, which merely doubles their age-specific cancer risk) the cancer onset rate is so low that impractically large numbers would be needed to characterize the moderate benefits that might realistically be hoped for. Even if there is a really worthwhile benefit from β -carotene, it may require 5 years of study of many thousands of people aged over 50 years before clear evidence of benefit emerges, so that complicated follow-up and any other source of organizational complexity will have to be avoided if such trials are to be maintained for several years.

Conclusion

Blood retinol: In malnourished populations, increasing the dietary intake of β -carotene or retinol among people deficient in vitamin A can rectify the deficiency and cause blood retinol levels to rise, a procedure of such obvious medical value that it matters little from a public health point of view whether it also diminishes cancer incidence. Although this may likewise hold for that small minority of people in developed countries whose diet is seriously deficient in vitamin A, for most people in developed countries neither dietary β -carotene ('provitamin A') nor dietary retinol ('preformed vitamin A') have much effect on blood retinol. If the inverse association between blood retinol and cancer incidence that has recently been reported from small prospective studies in developed countries is confirmed in larger prospective studies it may well reflect a true cause-and-effect relationship. Even if it does, however, this will not suggest that increasing dietary vitamin A or provitamin A will reduce cancer incidence in a developed country for they have little effect on blood retinol. Rather it would suggest that the hormonal and other extrinsic factors which do influence blood retinol levels need to be discovered, and that when they have been the effects on total health of manipulating them will need to be evaluated rigorously in long-term trials.

β -carotene: When dietary β -carotene rather than blood retinol is considered, it appears that so many observational studies of dietary factors and cancer incidence have already indicated a slightly lower than average incidence of cancer among people with an above average intake of β -carotene that it is most unlikely that this association will disappear entirely with future observational studies, whether of blood or of dietary β -carotene. This inverse association may, of course, be an artefact, due merely to association of β -carotene ingestion with some truly protective dietary habit(s) or component(s), or to association of β -carotene ingestion with avoidance of some truly harmful dietary habit(s) or component(s). Equally, however, the inverse association between dietary β -carotene and cancer incidence may be due to a genuine protective effect of β -carotene against the onset of cancer. If so, perhaps it is the β -carotene that is absorbed unchanged from the intestine which is chiefly responsible for this protective effect rather than any retinol or retinal generated in the intestine from β -carotene, in which case this protective effect could not be produced by dietary retinol alone. Although various lines of laboratory and field research can be envisaged which might tip the balance of probability a little one way or the other as to whether β -carotene is or is not truly protective, it seems unlikely that this question will be decided unequivocally in the foreseeable future unless controlled trials are undertaken. Although the evidence thus far available is not compelling that β -carotene is truly protective against cancer (or, *a fortiori*, against total mortality) this is not a good reason to delay starting controlled trials because even if, as seems probable, a few truly protective agents do await discovery among the dozen or two dietary factors of current interest to the research community, compelling evidence may take decades to emerge without controlled trials.

Preventive measures, especially those which may be relevant over a long period to many millions of people, deserve peculiarly rigorous evaluation. Apart from reducing the amount consumed (or harmfulness) of tobacco, there are few lines of research or governmental action that offer any medium-term hope of achieving substantial reductions in total cancer incidence rates. Seeking extrinsic determinants of the blood levels of carrier-bound retinol is one, and investigating the anticancer effects of the carotenoids is another. In each case, whether or not a fruitful outcome is thought to be probable, the fact that it is at least plausible would justify pursuit of some of the suggested lines of research.

We thank Bruce Ames for suggesting the relevance of singlet oxygen and John Cairns for suggesting how to test this, Professor Saxon Graham and his co-workers Curtis Mettlin and Jim Marshall, Dr M. Mathews-Roth, Dr E. Bjelke, Dr R. Shekelle,

Dr R. Willson, Dr H. Cuckle, Professor B. Modan, Professor M. P. Vessey, Dr C. R. Kay and Dr F. E. Speizer, for permission to mention their unpublished results, and Mrs V. M. Godwin for repeatedly preparing the typescript.

1. Smith, J. E. & Goodman, D. S. *Fedn Proc.* **38**, 2504–2509 (1979).
2. Goodman, D. S. *Fedn Proc.* **38**, 2501–2503 (1979).
3. De Luca, L. M. et al. *Report of the International Vitamin A Consultative Group* (Nutrition Foundation, New York, 1978).
4. Olson, J. A. & Hayaishi, O. *Proc. natn. Acad. Sci. U.S.A.* **54**, 1364–1370 (1965).
5. Nieman, C. & Obbink, H. J. K. *Vita. Horm.* **12**, 69–99 (1954).
6. Crisp, A. H. & Stonehill, E. *Postgrad. med. J.* **43**, 721–725 (1967).
7. Robboy, M. S., Sato, A. S. & Schwabe, A. D. *Am. J. clin. Nutr.* **27**, 362–367 (American Society for Microbiology, Washington DC, 1980); *Oncology* (in the press).
8. Mathews-Roth, M. M. et al. *J. Am. med. Ass.* **228**, 1004–1008 (1974).
9. McLaren, D. S. in *Dermatology in General Medicine* 2nd edn. (eds T. B. Fitzpatrick, A. Z. Eisen, K. Wolff, I. M. Freedberg & K. F. Austen) (McGraw-Hill, New York, 1979).
10. Tlustohowicz, R. C. et al. *Research Report 56* (Storrs Agric. Experiment Station, University of Connecticut, 1979).
11. Varma, R. N. & Beaton, G. H. *Can. J. Physiol. Pharmac.* **50**, 1026–1037 (1972).
12. Vahlquist, A., Michaelsson, G. & Juhlin, L. *Acta derm-venereol.* **48**, 437–442 (1978).
13. Underwood, B. A., Siegel, H., Wiesell, R. C. & Dolinski, M. *Am. J. clin. Nutr.* **23**, 1037–1042 (1970).
14. Gal, I., Parkinson, C. & Craft, I. *Brit. med. J.* **2**, 436–438 (1971).
15. *Nutr. Rev.* **37**, 346–348 (1979).
16. Smith, F. R. & Goodman, D. S. *New Engl. J. med.* **294**, 805–808 (1971).
17. Boutwell R. K. (consulting reviewer) *Selected Abstracts on Vitamin A in Cancer Biology* (International Cancer Research Data Bank, NCI, Bethesda, Maryland, 1979).
18. Schroder, E. W. & Black, P. H. *J. natn. Cancer Inst.* **65**, 671–674 (1979).
19. Sporn, M. B. & Newton, D. L. *Fedn Proc.* **38**, 2528–2534 (1979).
20. Peto, R. in *Origins of Human Cancer* (eds H. H. Hiatt, J. D. Watson & J. D. Winsten) (Cold Spring Harbor, New York, 1977).
21. Peto, R. *Proc. R. Soc. Lond. B* **205**, 111–120 (1979).
22. Rettura, G. et al. *170th nat. Meet. Am. chem. Soc., Abstr.* **59**, (1975).
23. Felix, E. L., Cohen, M. H. & Loyd, B. C. *J. surg. Res.* **21**, 307–312 (1976).
24. Foote, C. S. in *Free Radicals in Biology* Vol. 2 (ed. Pryor, W.) 85–133 (Academic, New York, 1976).
25. Krinsky, N. I. in *The Carotenoids* (ed. Isler, O.) 669–716 (Birkhäuser, Basle, 1971).
26. Epstein, J. H. *Photochem. Photobiol.* **25**, 211–213 (1977).
27. Mathews-Roth, M. M. in *Proc. 12th int. Congr. Chemotherapists* (American Society for Microbiology, Washington DC, 1980); *Oncology* (in the press).
28. Seifter, E. & Rettura, G. *175th nat. Meet. Am. chem. Ass. Abstr.* **45**, (1978).
29. Dorogokupla, A. G. et al. *Zdravookhr. Kazakh.* **10**, 32–34 (1973).
30. Bagdon, R. E., Zbinden, G. & Studer, A. *Tox. appl. Pharmac.* **2**, 225–236 (1960).
31. Kark, J. D., Smith, A. H., Switzer, B. R. & Hames, C. G. *J. natn. Cancer Inst.* **66**, 7–16 (1981).
32. Wald, N. et al. *Lancet* **ii**, 813–815 (1980).
33. Wahi, P. N. et al. *Indian J. Path. Bact.* **5**, 10–16 (1962).
34. Ibrahim, J., Jafarey, N. A. & Zuberi, S. J. *Clin. Oncol.* **3**, 58–63 (1977).
35. Clifford, P. *Proc. R. Soc. Med.* **65**, 24–28 (1972).
36. Basu, T. K. et al. *Int. J. Vitam. Nutr. Res.* **44**, 14–18 (1974).
37. Basu, T. K. et al. *Br. J. Cancer* **33**, 119–121 (1976).
38. Atukorala, S. et al. *Br. J. Cancer* **40**, 927–931 (1979).
39. Capel, I. D. & Williams, D. C. *ICRS med. sci.* **7**, 361 (1979).
40. Abels, J. C., Gorham, A. T., Pack, G. T. & Rhoads, C. P. *J. clin. Invest.* **20**, 749–764 (1941).
41. Cohen, M. H. et al. *Cancer Lett.* **4**, 51–54 (1977).
42. Mettlin, C. & Graham, S. *Am. J. Epidemiol.* **110**, 255–263 (1979).
43. Hirayama, T. *Nutr. Cancer* **1**(3), 67–81 (1979).
44. Dagadu, M. & Gillman, J. *Lancet* **i**, 531–532 (1963).
45. MacLennan, R. et al. *Int. J. Cancer* **20**, 854–860 (1977).
46. Bjelke, E. *Int. J. Cancer* **15**, 561–565 (1975).
47. Bjelke, E. in *Aktuelle Ernährungsmedizin* Suppl. 10–17 (Thieme, Stuttgart, 1978).
48. Mantel, N. & Haenszel, W. M. *J. natn. Cancer Inst.* **22**, 719–748 (1958).
49. Smith, P. G. & Jick, H. *Cancer* **42**, 808–811 (1978).
50. Gregor, A., Lee, P. N., Roe, F. J. C. et al. *Nutr. Cancer* **2**(2), 93–97 (1980).
51. Mettlin, C., Graham, S. & Swanson, M. *J. natn. Cancer Inst.* **62**, 1435–1438 (1979).
52. Mettlin, C., Graham, S., Priore, R., Marshall, J. & Swanson, M. *Nutr. Cancer* (in the press).
53. Graham, S. et al. *Am. J. Epidemiol.* (in the press).
54. Marshall, J., Graham, S., Mettlin, C. & Swanson, M. *Abstr. 13th Meet. Soc. Epidemiologic Research*, Minneapolis (1980).
55. Haenszel, W., Kurihara, M. & Segi, M. *J. Natn. Cancer Inst.* **49**, 969–988 (1972).
56. Haenszel, W. et al. *J. natn. Cancer Inst.* **51**, 1765–1779 (1973).
57. Modan, B. et al. *Cancer* **34**, 2087–2092 (1974).
58. Modan, B. et al. *J. natn. Cancer Inst.* **55**, 15–18 (1975).
59. Modan, B., Cuckle, H. & Lubin, F. (in preparation).
60. Tuyns, A. J., Pequinot, G. & Jensen, O. M. *Bull. Cancer* **65**, 59–64 (1978).
61. Cook-Mozaffari, P. *Nutr. Cancer* **1**(2), 51–60 (1979).
62. Cook-Mozaffari, P. et al. *Br. J. Cancer* **39**, 293–409 (1979).
63. Stocks, P. in *British Empire Cancer Campaign 35th a. Rep. 1957*, Part II, Suppl., 111–113 (1958).
64. Phillips, R. L. *Cancer Res.* **35**, 3513–3522 (1975).
65. Paymaster, J. C., Sanghvi, L. D. & Gangadharan, P. *Cancer* **21**, 279–288 (1968).
66. Wynder, E. L. et al. *Cancer* **16**, 1461–1496 (1963).
67. Mathews-Roth, M. M., Crean, C. & Clancy, M. *Nutr. Rep. Int.* **16**, 419–423 (1977).
68. Mufson, R. A., DeFeo, D. & Weinstein, I. B. *Molec. Pharmac.* **16**, 569–578 (1979).
69. Foote, C. S., Chang, Y. C. & Denny, R. W. *J. Am. chem. Soc.* **92**, 5216–5218 (1970).
70. Packer, J. E. et al. *Biochem. biophys. Res. Commun.* **98**, 901–906 (1981).
71. Oberley, L. W. & Buettner, G. R. *Cancer Res.* **39**, 1141–1149 (1979).
72. Demopoulos, H. B., Pietronigro, J. D., Flamm, E. S. & Seligson, M. L. *J. envir. Path. Tox.* **3**, 273–303 (1980).
73. Ames, B. N., Holstein, M. C. & Cathcart, R. in *Lipid Peroxidation in Biology and Medicine* (ed. Yagi, K.) (Academic, New York, in the press).
74. Reddy, J. K., Azarnoff, D. L. & Hignite, C. E. *Nature* **283**, 397–399 (1980).
75. Shamberger, R. J. *J. natn. Cancer Inst.* **48**, 1491–1494 (1972).
76. Patwardhan, V. N. *Am. J. clin. Nutr.* **22**, 1106–1118 (1969).
77. Leitner, Z. A., Moore, T. & Sharman, I. M. *Proc. Nutr. Soc.* **34**, 44A (1975).
78. *Wd H th Org. tech. rep. Ser.* no. 557 (1974); *FAO Nutr. Meet. Rep. Ser.*, 1974, no. 54 (FAO, Rome, 1975).
79. Mathews-Roth, M. M. et al. *New Engl. J. Med.* **282**, 1231–1234 (1970).
80. Goerz, G. & Ippen, H. *Di. med. Wsch.* **102**, 1051–1055 (1977).
81. Zaynoun, S. T. et al. *Brit. J. Derm.* **97**, 663–668 (1977).
82. Vessey, M. P. et al. *Lancet* **i**, 549–550 (1981).
83. Laydale, P. M. et al. *Lancet* **i**, 541–546 (1981).
84. Chope, H. D. & Breslow, L. *Am. J. publ. Hlth* **46**, 61–67 (1956).

ARTICLES

Earth's neodymium budget and structure and evolution of the mantle

Geoffrey F. Davies

Department of Earth and Planetary Sciences, McDonnell Center for the Space Sciences, Washington University, St Louis, Missouri 63130, USA

Neodymium and strontium isotopic data from rocks can be satisfied by a model in which the mantle contains chemical heterogeneities of many sizes which are less depleted in incompatible elements than the surrounding mantle, and in which the entire mantle is slowly mixed by convection.

NEODYMIUM isotopic abundance data have been used to constrain models of the structure and evolution of the Earth's mantle^{1–5}. According to some studies, only a fraction (one-third to one-half^{1,3–5}) of the mantle has been differentiated to form the continental crust, and the rest of the mantle is thought to retain its primitive Nd isotopic composition. This would favour models in which the mantle is compositionally layered, but not those in which the mantle is mixed on the large scale by whole-mantle convection. As our ideas on the evolution of the Earth will be strongly affected by which model is correct (geophysical constraints having proved inconclusive^{6–11}) so we need to establish the uncertainties of the isotopic constraints and the interpretations. Unfortunately, recent studies have only briefly considered the uncertainties of the constraints, they have assumed that only layered mantle models are viable, and the

complexities of calculations of the evolution of these models have obscured their sensitivity to uncertainties in the data.

This article examines the uncertainties of the Nd isotopic constraints; this is simplified by considering the present Nd budget of the Earth separately from the evolution of the Earth and of Nd. The mass of mantle which has been depleted to form the continental crust is found to be uncertain, and identifying this with the seismologically defined upper mantle is premature.

Some deficiencies of layered mantle models are also pointed out—the narrow range of Nd isotope data for ocean island basalts is not explained, and an important part of the rationale for layered models may be undermined by recent Nd data for kimberlite inclusions.

A model with a heterogeneous mantle and whole-mantle convection is proposed. It is pointed out that whole-mantle

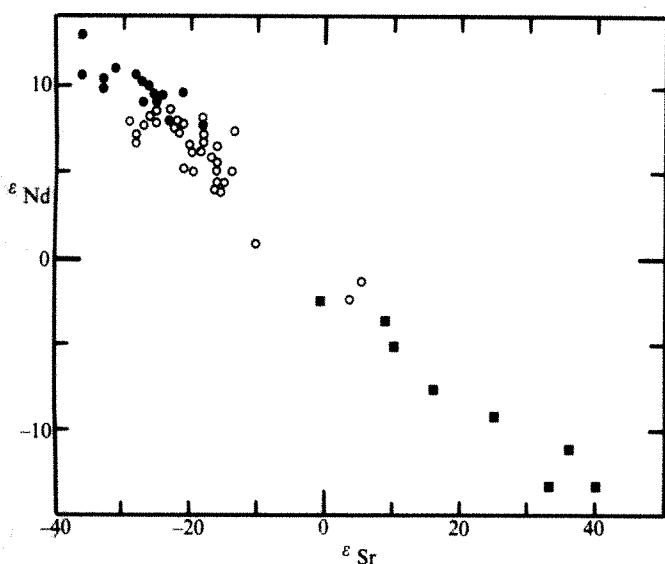


Fig. 1 Nd-Sr isotopic correlation for some oceanic rocks and kimberlite inclusions. Mid-ocean ridge basalt (MORB, ●) and oceanic island basalt (OIB, ○) data from ref. 2. Kimberlite inclusion (KI, ■) data from ref. 22. $\epsilon_{Nd} = 10^4(\alpha_{Nd}/0.511836 - 1)$ (refs 4, 12, 13). $\epsilon_{Sr} = 10^4(\alpha_{Sr}/0.7045 - 1)$ (refs 12, 14), where $\alpha_{Sr} = {}^{87}\text{Sr}/{}^{86}\text{Sr}$.

convection involves long overturn times and very long residence times, that plate tectonics must continuously generate heterogeneities, and that both mechanical and thermochemical factors may extend the survival times of heterogeneities to thousands of millions of years.

Nd isotope data

Variations in $\alpha_{Nd} = {}^{143}\text{Nd}/{}^{144}\text{Nd}$ produced by the decay of ${}^{147}\text{Sm}$ are conveniently measured in terms of ϵ_{Nd} (ref. 12), which measures parts in 10^4 deviations of α_{Nd} from a reference value that is based on measurements of the Juvinas meteorite¹³, and is assumed to represent an average chondritic value. Mid-ocean ridge basalts (MORB) have ϵ_{Nd} values near 10 (refs 12, 14–16), indicating that the MORB source (the oceanic mantle) is depleted in large-ion lithophile (LIL) elements relative to chondrites. Conversely, many young continental rocks have negative ϵ_{Nd} values^{17–20} indicating a complementary LIL-enrichment relative to chondrites. Island arc rocks tend to have $\epsilon_{Nd} \sim 8$, while oceanic island basalts (OIB) tend to have $\epsilon_{Nd} \sim 6$, but with some ranging to near zero^{2,12,14,16}. Continental flood basalts (CFB) tend to have ϵ_{Nd} values near zero²¹, which suggests they derive from a source which is primitive and chondritic^{1,4}. A negative correlation between ϵ_{Nd} and ϵ_{Sr} has been found for oceanic rocks^{12,14–16}, where ϵ_{Sr} is the analogous measure of the Rb-Sr system¹² with $\epsilon_{Sr} = 0$ being defined by the point where the correlation crosses $\epsilon_{Nd} = 0$. This correlation is shown in Fig. 1 for a selection of MORB and OIB samples². Recently some kimberlite inclusion (KI) data have been obtained which extend this correlation through the origin²² (Fig. 1). Other continental rocks and island arc rocks often deviate from this correlation, and this can often be attributed to contamination by seawater or lower continental rocks of magma originally lying on the correlation line^{17,20,23–25}.

Nd in chondrites

Interpretations of the above data have been based on the assumptions that the average crust-mantle system (bulk Earth or BE) has chondritic Sm/Nd and Rb/Sr ratios, and that chondritic material will have $\epsilon_{Nd} = 0$. The first assumption is very difficult to justify with any rigour. For example, if chondrites derive from a relatively small number of parent bodies, then our sample may be significantly biased. Furthermore, the Earth may have sampled a slightly different population of material, because

many meteorites are believed to be derived from the asteroid belt, more than twice as far from the Sun as the Earth.

There are some limited data with which the second assumption has been tested. Jacobsen and Wasserburg²⁶ have reported Sm and Nd isotopic and abundance data for several chondrites and achondrites (including Juvinas). Their results imply $\epsilon_{Nd} = -1 \pm 1$ at present and $\epsilon_{Nd} = -2 \pm 0.5$ at 4,500 Myr ago, but it is not clear to what extent these discrepancies from $\epsilon_{Nd} = 0$ may be due to interlaboratory biases^{26,27}. (Note that these ϵ values are based on the old reference values^{12,13}, and not on the new ones defined by Jacobsen and Wasserburg²⁶.) About 50 isotope dilution measurements of Sm/Nd in chondrites^{28–32} give a mean ${}^{147}\text{Sm}/{}^{144}\text{Nd}$ of 0.196 ± 0.002 (see also refs 26, 27), compared with the Juvinas reference value of 0.1936 (ref. 13). If all chondritic material began with $\epsilon_{Nd} = 0$, this Sm/Nd range would generate a range of 2 ± 6 (1 s.d.) in ϵ_{Nd} , with the mean being $\epsilon_{Nd} = 2 \pm 2$ (2 s.d.).

Present Nd mass balance

Suppose that the mantle initially had mass M , Nd mass concentration C_0 and $\epsilon_{Nd} = \epsilon_0$. Suppose that later the mantle comprises several reservoirs with masses M_i , concentrations C_i and $\epsilon_{Nd} = \epsilon_i$, and the mass fraction $x_i = M_i/M$. Then the equations which express conservation of the masses of the mantle and Nd and the total amount of radiogenic Nd produced in the interval are

$$\begin{aligned} \sum x_i &= 1 \\ \sum x_i C_i &= C_0 \\ \sum x_i C_i \epsilon_i &= C_0 \epsilon_0 \end{aligned} \quad (1)$$

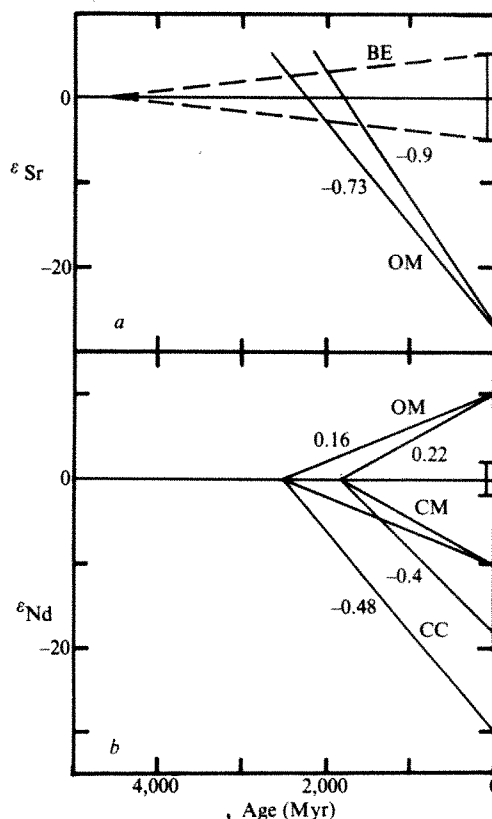


Fig. 2 Models of ϵ_{Sr} and ϵ_{Nd} evolution calculated from equation (1) and its Sr analogue ($Q_{Sr} = 0.0167 \text{ Myr}^{-1}$, ref. 4). *a*, Evolution of ϵ_{Sr} in the oceanic mantle (OM) relative to the bulk Earth (BE). OM lines are labelled with the Rb-Sr enrichment factor, f_{Sr} . Error bar on BE is the estimated uncertainty of the present BE value of ϵ_{Sr} (see text). $f_{Sr} = (\mu_{Sr}/0.0839 - 1)$, where $\mu_{Sr} = {}^{87}\text{Rb}/{}^{86}\text{Sr}$ (refs 4, 12, 14). *b*, Evolution of ϵ_{Nd} for oceanic mantle (OM), continental mantle (CM) and continental crust (CC) for mean differentiation ages of 1,800 and 2,500 Myr. Lines are labelled with values of f_{Nd} .

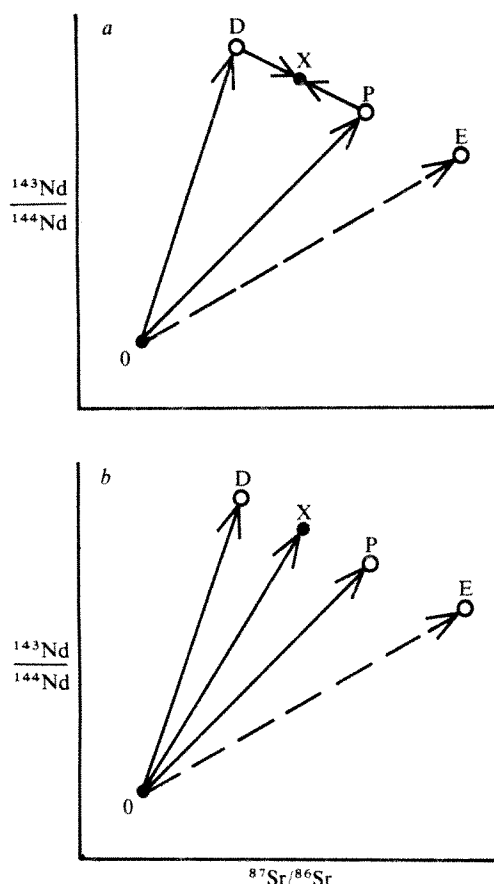


Fig. 3 Alternative explanations of the correlation in Fig. 2. *a*, Depleted (D) and primitive (P) sources evolve independently from 0. Mixing from D and P then yields the isotopic composition X. The presence of an enriched source (E), as suggested by kimberlite inclusion data²², implies that either E evolved independently of but colinearly with D and P, or that P results from mixing of D and E and is therefore not primitive. Straight mixing lines require the sources to have the same Sr/Nd ratio⁴⁶. *b*, All sources evolve independently but colinearly. This requires coherent fractionation of Sm/Nd and Rb/Sr: specifically, that sources have the same ratio of enrichment factors, f_{Nd}/f_{Sr} (refs 2, 53).

These relations are implicit in the treatments of Nd evolution and partitioning by O'Nions *et al.*³ and Jacobsen and Wasserburg⁴, but they did not separate the instantaneous mass balance from the evolution. In equations (1), the evolution is implicit in the present values of ϵ .

Some results of applying equations (1) are given in Table 1. Cases 1–4 are variations on Model 1 of Jacobsen and Wasserburg⁴ (which is similar to Case 3, Table 1): the present reservoirs considered are the continental crust (CC), depleted mantle (DM), and primitive mantle (the same as the initial mantle, but with a smaller mass). The parameters of the initial mantle (IM) and of CC were assumed, and those of DM were calculated, assuming $\epsilon_{DM} = 10$, consistent with Fig. 1 (ref. 4). The initial concentration of Nd, $C_0 = 0.9 \mu\text{g per g}$, is lower than that assumed by Jacobsen and Wasserburg⁴, and is based on thermal evolution models of the Earth^{33,34} which require uranium abundances in the mantle to be ~ 1.5 times chondritic: C_0 is also assumed to be 1.5 times chondritic, rather than twice chondritic⁴.

Case 1, Table 1, requires about half of the mantle to have been depleted ($x_{DM} = 0.48$). If ϵ_{CC} is assumed to be -30 rather than -20 , which implies an older crust, then about two-thirds of the mantle must have been depleted (Case 2). On the other hand, if $C_0 = 1.26 \mu\text{g per g}$ (ref. 4), then only one-third of the mantle need have been depleted (Case 3). The sensitivity of the mass

balance to ϵ_0 is illustrated by Case 4: almost the entire mantle must have been depleted if ϵ_0 is increased to $+4$.

There is increasing evidence that the sub-continental mantle is veined by LIL-element-enriched phases which were emplaced by metasomatic events into previously depleted mantle^{23,35–37}. The kimberlite-inclusion data in Fig. 1 suggest that $\epsilon_{Nd} = -10$ may be reasonably representative of this heterogeneous material. Its total abundance is difficult to estimate, but one constraint may come from continental heat flow data, from which it is inferred that $\sim 25 \text{ mW m}^{-2}$ of heat flux comes from below the upper 10–15 km of the crust³⁸: this is compatible with uranium and thorium abundances in the continental lithosphere up to ~ 10 times chondritic abundances (that is, U abundances up to $\sim 0.15 \mu\text{g per g}$). Jordan³⁹ points out that continental garnet lherzolite inclusions have K and other LIL abundances compatible with this degree of enrichment, and he summarizes the confluence of geophysical and geological data favouring the existence of an ancient, stable, LIL-enriched subcontinental mantle. Thus the subcontinental mantle may be an ancient reservoir with significant Sm and Nd enrichments.

Cases 5–8 of Table 1 explore the possible effect of an additional reservoir in the sub-continental lithospheric mantle (CM) by giving the Nd concentration in CM which would be required by depletion of the entire remaining mantle ($x_{CM} = 0.03$ corresponds to continental lithosphere $\sim 200 \text{ km}$ thick). Cases 5, 6 and 8 yield abundances from ~ 3 – 13 times chondritic ($0.6 \mu\text{g per g}$); only Case 7 clearly violates the constraint inferred from heat flow.

These results demonstrate that the fraction of mantle which has remained undepleted is subject to large uncertainties. Thus, the identification of the undepleted mantle with the lower mantle is speculative, and most of the mantle may possibly have been depleted.

Age of the continental crust

Because of the long half life of ^{147}Sm , ϵ_{Nd} varies approximately linearly with time in a closed Sm–Nd system, and can be written⁴:

$$\epsilon_{Nd}(t) = \epsilon_{Nd}(0) + Q_{Nd}f_{Nd}t \quad (2)$$

where $Q_{Nd} = 0.02474 \text{ Myr}^{-1}$ is a constant related to the decay constant, t is time and f_{Nd} is the Sm–Nd enrichment factor of the reservoir relative to Juvinas: $f_{Nd} = (\mu_{Nd}/0.1936 - 1)$ and $\mu_{Nd} = ^{147}\text{Sm}/^{144}\text{Nd}$. Jacobsen and Wasserburg⁴ have shown that equation (2) yields an upper bound, T , on the mean age of a reservoir if $\epsilon(-T)$, $\epsilon(0)$ and f_{Nd} are known.

The chronology of the continental crust is a major question in the mass-balance calculations (and in Earth history in general). Equation (2) shows explicitly how the evolution of reservoirs enters the mass balance calculations of the previous section. For example, Case 2 (Table 1), with $\epsilon_{CC} = -30$, would require the continental crust to be older, or to have a more negative f_{Nd} , or both, than Case 1, with $\epsilon_{CC} = -20$.

Table 1 Nd mass balance

Case	Initial mantle ($x = 1$) C ($\mu\text{g per g}$)	ϵ	x	Depleted mantle ($\epsilon = 10$) C ($\mu\text{g per g}$)	Continental crust ($x = 0.0055$, $C = 26$) ϵ	Continental mantle ($x = 0.03$, $\epsilon = -10$) C ($\mu\text{g per g}$)
1	0.9	0	0.48	0.60	-20	—
2	0.9	0	0.64	0.68	-30	—
3	1.26	0	0.34	0.84	-20	—
4	0.9	4	0.93	0.75	-25	—
5	0.9	0	0.97*	0.53	-20	8
6	0.9	0	0.97*	0.60	-30	5
7	1.26	0	0.97*	0.74	-20	14
8	0.9	4	0.97*	0.71	-20	1.6

* Assumed.

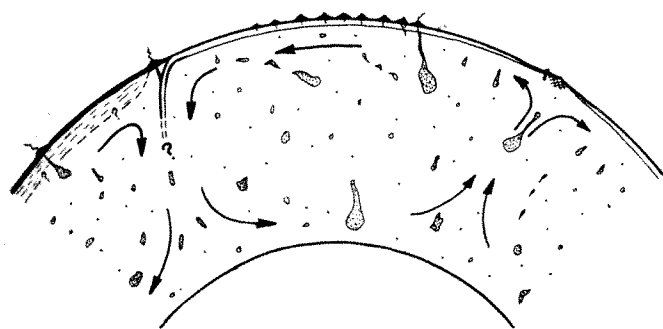


Fig. 4 Some aspects of the heterogeneous mantle models. Blobs represent chemical heterogeneities which are less LIL-depleted than the surrounding mantle. Heterogeneities are of many sizes, but their geometry is unknown; they may be generated at subduction zones and higher concentrations of radioactive heat sources would make them prone to diapirism; they would tend to be destroyed or dispersed by the shearing of convective flow.

Jacobsen and Wasserburg⁴ estimated the mean age of the crust-mantle differentiation using the Rb-Sr system in the oceanic mantle (OM). For OM, the present ϵ_{Sr} is about -27 (Fig. 1), and they estimate that f_{Sr} lies between -0.8 and -0.95, with -0.9 being their preferred value. The mean age, T_{Sr} , of OM can then be obtained from the Sr analogue of equation (2), and the graphical solution is shown in Fig. 2a: assuming that initially OM had $\epsilon_{Sr} = 0$, the result is $T_{Sr} \leq 1,800$ Myr. This is a rather small age for the depleted OM and its complementary continental crust (CC), and would favour Cases 1 and 3 ($\epsilon_{Nd} = -20$) rather than Case 2 ($\epsilon_{Nd} = -30$), as is shown in Fig. 2b assuming that $f_{Nd} = -0.4$ for CC⁴.

Jacobsen and Wasserburg's^{4,40} confidence in this estimate is based on the presumption that the OM is strongly depleted in Rb, so that f_{Sr} must be close to -1 and its absolute uncertainty cannot be large. However, Weaver and Tarney⁴¹, for example, estimate $f_{Sr} = -0.73$ for OM, which yields $T_{Sr} \leq 2,200$ Myr (Fig. 2a). Note that the bulk Earth value of ϵ_{Sr} is inferred from the correlation in Fig. 1 and is, therefore, uncertain by about ± 5 (and the correlation may not even preserve direct information about the initial ϵ_{Sr} of the Earth). When account is taken of such uncertainties in enrichment factors and ϵ values of various reservoirs^{3,4,41}, one must conclude that equation (2) and its Sr analogue do not yet give very strong constraints on the mean age of the crust-mantle differentiation (Fig. 2a).

On the other hand, the observations that the continents contain provinces which were stabilized by the late Archaean (~2,500 Myr ago), are of average crustal thickness (35-40 km) and which have little or no sediment cover lead to the conclusion that much of the continental crust had been differentiated by the late Archaean (see refs 42-44). Wise⁴² estimates that 90% of the continental crust had been differentiated by the end of the Archaean. This conclusion can be avoided if the Archaean crust had subsequently been thickened by underplating, but there is no recognized modern analogue of such a process, and it would have to be an extremely delicate process to affect broad areas uniformly without leaving recognizable effects at the surface. This conclusion would imply that the mean age of differentiation of the continental crust from the mantle is $\geq 2,500$ Myr.

If the mean age of the continental crust is as great as 2,500 Myr, and with estimates of f_{Nd} of CC ranging from -0.35 to -0.48 (refs 3, 4, 41, 45), then it is quite plausible for ϵ_{Nd} of CC to be -30, as assumed in Table 1. It follows that well over half of the mantle may be LIL-depleted. Note that a Sr mass balance may also be calculated^{3,4} as was done here for Nd, but the uncertainties for that calculation are at least as great: the mean Rb-Sr enrichment of the continental crust, in particular, is very uncertain.

The layered mantle model

Wasserburg and DePaolo¹ have advocated a two-reservoir mantle model in which the (seismologically defined) upper mantle is identified as depleted oceanic mantle and the lower mantle is identified as a primitive (chondritic) source of the continental flood basalts. This division is suggested by their mass balance estimates^{1,4} (see case 3, Table 1). In this model the correlation in the depleted quadrant of Fig. 1 comprising MORB, island arcs, OIB and CFB (not shown, but with $\epsilon_{Nd} \approx 0$), is interpreted as a mixing line between MORB and CFB. If that is true, then the clustering of the OIB data at $\epsilon_{Nd} = 6 \pm 2$ is puzzling: why do not all degrees of mixing occur? As no one seems to advocate a special intermediate source reservoir with $\epsilon_{Nd} = 6$, the clustering must be due either to the small data sample or to some unknown mechanism. If either of these is true for the OIB, then why not also for the CFB, clustering at $\epsilon_{Nd} = 0$, especially because the new KI data suggest that the correlation of Fig. 1 extends into the enriched quadrant?

There are two mechanisms which might yield a correlation like that in Fig. 1, and these are shown in Fig. 3: mixing between two reservoirs (Fig. 3a), and coherent fractionation with independent evolution of sources (Fig. 3b). As long as only two reservoirs were sufficient to account for the data, the mixing hypothesis was perhaps preferable. However, the advent of the KI data (Fig. 1) necessitates at least one more reservoir: an enriched KI endmember (E, Fig. 3). If these reservoirs have evolved independently and colinearly, then coherent fractionation is established and can be invoked for any other point on the correlation line. Alternatively, if all points between D and E result from mixing of D and E (requiring identical Sr/Nd (ref. 46)), then it would only be coincidence that the CFB have $\epsilon_{Nd} = 0$, the identification of the CFB source as primitive would be premature, and the primitive values of ϵ_{Sr} would be unknown. Furthermore, the geometry of the sources would have to be revised. In either case, the KI data make the case for the layered two-reservoir mantle model much less compelling.

Another important question about the layered mantle model is how the layers separate (note that the separation must not be complete, as lower mantle material is required to reach the surface). The mantle transition zone may effect this separation, either by the effects of phase changes or, more likely, by a change in intrinsic density⁶. If the upper mantle is separated in this way and has been steadily depleted, then, according to Jacobsen and Wasserburg⁴, the concentrations of some highly incompatible elements in crustal rocks would have decreased 10-fold in the past 3,500 Myr, and they claim there is no evidence for this. If, as they prefer, the depleted mantle has been steadily growing in volume, with a constant degree of depletion, then it would only be coincidence if its lower boundary coincided with the transition zone at present. If the boundary between depleted and undepleted layers does not coincide, or has not coincided, with the transition zone, then the existence of the boundary and its ability to prevent mixing become *ad hoc*: it is certainly possible, but there is no compelling geophysical evidence for another boundary in the mantle.

Heterogeneous mantle model

At this stage, the Nd data can be interpreted in terms of an alternative model of the mantle which is chemically heterogeneous up to the largest scales and in which the entire mantle is slowly mixed by convection. The geophysical evidence is equivocal on the question of the structure of mantle convection: both whole-mantle convection^{6,7,11} and convection in separate layers⁸⁻¹⁰ have been recently advocated on the basis of geophysical evidence, so geochemists must also consider these alternatives.

The type of model proposed is shown in Fig. 4. It is postulated that the average mantle is fairly strongly LIL-depleted ($\epsilon_{Nd} \approx +10$) and that a small volume fraction (perhaps 10-20%) comprises less depleted or enriched regions. These regions are depicted in Fig. 4 as blobs, but their geometry is unknown. They

could be remnants of primitive mantle ($\epsilon_{Nd} = 0$), but it is more likely that they are products of more complex differentiation histories involving both enrichments and depletions, resulting in a wider range of ϵ_{Nd} ($10 \geq \epsilon_{Nd} \geq -10$). If these differentiation events involved coherent fractionation, as at least some must have, according to the earlier discussion, then the linear array of Fig. 1 would result. The heterogeneities may range in size up to hundreds or even thousands of kilometres, as suggested by the largest oceanic volcanic hotspots (Iceland and Hawaii), and down to grain sizes³⁵⁻³⁷. The mantle heterogeneities must have existed through much of the Earth's history, as indicated by the mean age of the crust and mantle reservoirs (Fig. 2), and by inferred ancient metasomatic events²³.

Subduction zones must generate mantle heterogeneities, because subducting oceanic lithosphere is strongly stratified and it apparently remains mechanically coherent at least to depths of hundreds of kilometres. The upper mantle part of subducted lithosphere is presumably LIL-depleted and refractory: it may therefore be mechanically strong and persistent, while pieces of it may ultimately become incorporated into continental lithosphere^{39,47}. The crustal part is enriched in some LIL elements and hydrated⁴⁰; much of the enriched component may be incorporated into island arc magmas^{20,24,25}, but some probably escapes this process and returns to the mantle⁴³. Until the complexities of island arc geochemistry and of sialic crust generation are better understood, it is difficult to predict the character of such LIL mantle enrichments, but if they are significantly enriched in radioactive heat sources (U, Th, K), they would be prone to diapiric instability and uprising as mantle viscosity is a strong function of temperature. Such diapirs would be likely to mobilize and concentrate volatile and other incompatible elements, and thus might be the cause of mantle metasomatic events. They would be, in this model, the immediate sources of volcanic hotspots, although only a small fraction of all existing diapirs might reach the surface at a given time; many others might be disrupted by shearing before reaching the surface. Some of these features are shown in Fig. 4.

An important question is how the postulated heterogeneities could survive in a fully convective mantle, as the shearing in convective flow tends to smear out heterogeneities^{48,49}: the following arguments demonstrate the plausibility of heterogeneities surviving for thousands of millions of years. In a whole-mantle convection system, at present rates of seafloor spreading, it would take over 4,000 Myr for the mass of the mantle to pass within 100 km of the surface⁵⁰, and thus to undergo a significant melting and differentiation event; that is the mean residence time between major differentiation events is thousands of millions of years. For isolated more-viscous regions, shearing rates vary inversely as viscosity, so that refractory regions are likely to have low shearing rates. On the other hand, shearing rates of isolated less viscous regions are controlled by the viscosity of the surrounding mantle and thus will not be exceptionally large. In addition, the diapiric tendency of such regions, suggested above, will tend to enhance enriched heterogeneities.

A heterogeneous mantle is a natural outcome of the present tectonic regime, but the degree of chemical heterogeneity is difficult to estimate as it will be the result of a dynamic equilibrium

between the competing processes creating and destroying heterogeneities. This can be characterized in terms of the residence times of mantle material and survival times of heterogeneities. Residence times are clearly of the order of 10^9 years at present. Survival times depend on the rates of competing processes and these are difficult to estimate. The shearing rate can be characterized by the mantle overturn time which, at 30 mm yr^{-1} , is at least 300 Myr for the whole mantle. In a homogeneous fluid, heterogeneities can become very elongated by shearing in one overturn^{46,47}, but the mechanisms discussed above might easily extend the survival time well beyond one overturn time.

Because the Earth's heat flow was probably greater in the past, all processes would have been faster. Plate velocities are approximately proportional to the square of the heat flux⁵¹, but it is unlikely that the heat flux has been much more than twice its present value in the past 2,000 Myr (refs 33, 52). Thus residence times have probably been $> 1,000$ Myr for the past 2,000 Myr, and other rates would probably have changed in proportion.

O'Connell and Hager⁵⁰ have pointed out that residence times may be considerably less in the upper mantle than in the lower mantle because of the relatively rapid migration of ridges and trenches compared with convective velocities. Essentially, the more recently subducted material will have a greater chance of being returned to the surface by a passing ridge than deeper material. Thus this kinematic effect would tend to produce a stratification (but not necessarily layering) of the mantle in which the deeper material has been less differentiated than shallower material.

Conclusions

The Nd and Sr isotopic data clearly demand a mantle in which heterogeneities have existed for thousands of millions of years: the real question is the topology and distribution of the heterogeneities. In this sense the layered mantle model is a special case in which there are only vertical heterogeneities with sharp boundaries. Our heterogeneous mantle model is another type in which there is little or no mean vertical variation and in which the heterogeneities are isolated and have a range of chemical differences and a large range of sizes. The principal differences between these models lie in (1) the fraction of the mantle which has been strongly LIL-depleted, (2) the topology of the heterogeneities and (3) the existence of a reservoir of isotopically primitive material.

The principal question about the heterogeneous mantle model is probably the uncertainty of the survival time of the heterogeneities, while the principal question about the layered model may be the mechanism by which the layers are separated.

If some heterogeneities must exist within the layers of the layered mantle model, and if the heterogeneous model has some vertical stratification, then the distinctions between the models begin to blur. The real issue then is not a choice between two idealized models, but the accuracy with which the Earth's place in the spectrum of models can be determined. It is concluded here that no part of the spectrum of models discussed above can confidently be excluded at present.

I thank Don DePaolo, Gerry Wasserburg, Rodey Batiza, Larry Haskin and Keith O'Nions for constructive comments.

Received 31 July 1980; accepted 5 January 1981.

1. Wasserburg, G. J. & DePaolo, D. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3594-3598 (1979).
2. Allegre, C. J., Ben Othman, D., Polve, M. & Richard, P. *Phys. Earth planet. Inter.* **19**, 293-306 (1979).
3. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *J. geophys. Res.* **84**, 6091-6101 (1979).
4. Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7411-7427 (1979).
5. DePaolo, D. J. *Geochim. cosmochim. Acta* **44**, 1185-1196 (1980).
6. Davies, G. F. *Geophysical J. R. astr. Soc.* **49**, 459-486 (1977).
7. Hager, B. H. & O'Connell, R. J. *J. geophys. Res.* **84**, 1031-1048 (1979).
8. Richter, F. M. *J. geophys. Res.* **84**, 6783-6795 (1979).
9. Parmentier, E. M. & Turcotte, D. L. *Phys. Earth planet. Inter.* **17**, 281-289 (1978).
10. McKenzie, D. P. & Weiss, N. O. *Geophys. J. R. astr. Soc.* **42**, 131-174 (1975).
11. Jordan, T. H. *J. Geophys.* **43**, 473-496 (1977).
12. DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 249-252, 743-746 (1976); **4**, 465-468 (1977).
13. Lugmair, G. W., Scheinin, N. B. & Marti, K. *Earth planet. Sci. Lett.* **27**, 79-84 (1975).
14. DePaolo, D. J. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **43**, 999-1008 (1979).

15. O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **34**, 13-22 (1977).
16. Richard, P., Shimizu, N. & Allegre, C. J. *Earth planet. Sci. Lett.* **31**, 269-278 (1976).
17. Carter, S. R., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Science* **202**, 743-747 (1978).
18. Allegre, C. J. & Ben Othman, D. *Nature* **286**, 335-342 (1980).
19. McCulloch, M. T. & Wasserburg, G. J. *Science* **200**, 1003-1011 (1978).
20. Hawkesworth, C. J. & Vollmer, R. *Contr. Miner. Petrol.* **69**, 151-165 (1979).
21. DePaolo, D. J. & Wasserburg, G. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3056-3060 (1979).
22. Menzies, M. & Murthy, V. R. *Nature* **283**, 634-636 (1980).
23. Menzies, M. & Murthy, V. R. *Earth planet. Sci. Lett.* **46**, 323-334 (1980).
24. DePaolo, D. J. *Science* **209**, 684-687 (1980).
25. Hawkesworth, C. J. *et al. Earth planet. Sci. Lett.* **42**, 45-57 (1979).
26. Jacobsen, S. B. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **50**, 139-155 (1980).
27. Lugmair, G. W. & Carlson, R. W. *Proc. 9th Lunar planet. Sci. conf.* 689-704 (1978).
28. Masuda, A., Nakamura, N. & Tanaka, T. *Geochim. cosmochim. Acta* **37**, 239-248 (1973).
29. Nakamura, N. *Geochim. cosmochim. Acta* **38**, 757-775 (1974).
30. Nakamura, N. & Masuda, A. *Earth planet. Sci. Lett.* **19**, 429-437 (1973).
31. Gast, P. W., Hubbard, N. J. & Wiesmann, H. *Proc. 2nd Lunar Sci. Conf.*, 1143-1163 (1970).

32. Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Geochim. cosmochim. Acta* **42**, 1199–1212 (1978).
33. Davies, G. F. *J. geophys. Res.* **85**, 2517–2530 (1980).
34. Schubert, G., Stevenson, D. & Cassen, P. *J. geophys. Res.* **85**, 2531–2538 (1980).
35. Lloyd, F. E. & Bailey, D. K. *Phys. Chem. Earth* **9**, 385–416 (1975).
36. Frey, F. A., Green, D. H. & Roy, S. D. *J. Petrol.* **19**, 463–513 (1978).
37. Frey, F. A. & Green, D. H. *Geochim. cosmochim. Acta* **38**, 1023–1059 (1974).
38. Sclater, J. G., Jaupart, C. & Galson, D. *Rev. Geophys. Space Phys.* **18**, 269–312 (1980).
39. Jordan, T. H. *Phil. Trans. R. Soc.* (in the press).
40. Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7429–7445 (1979).
41. Weaver, B. L. & Tarney, J. *Nature* **286**, 342–346 (1980).
42. Wise, D. U., in *The Geology of Continental Margins* (eds Burke, C. A. & Drake, C. L.) 45–58 (Springer, New York, 1974).
43. Armstrong, R. L. *Rev. Geophys. Space Phys.* **6**, 175–200 (1968).
44. Hess, H. H. in *Petrologic Studies* (eds Engel, A. E. J., James, H. L. & Leonard, B. F.) 599–620 (Geological Society of America, New York, 1962).
45. DePaolo, D. P. *Geochim. cosmochim. Acta* **44**, 1185–1196 (1980).
46. DePaolo, D. J. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **43**, 615–627 (1979).
47. Oxburgh, E. R. & Parmentier, E. M. *Phil. Trans. R. Soc. A* **288**, 415–428 (1978).
48. McKenzie, D. P. *Geophys. J. R. astr. Soc.* **58**, 689–715 (1979).
49. Richter, F. M. & Ribe, N. M. *Earth planet. Sci. Lett.* **43**, 212–222 (1979).
50. O'Connell, R. J. & Hager, B. H. *EOS* **61**, 373 (1980).
51. Davies, G. F. *Earth planet. Sci. Lett.* **44**, 231–238 (1979).
52. Davies, G. F. *J. geophys. Res.* **85**, 7108–7114 (1980).
53. DePaolo, D. J. *Earth planet. Sci. Lett.* **43**, 201–211 (1979).

Structure of the neuraminidase gene in human influenza virus A/PR/8/34

Stan Fields & Greg Winter

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

George G. Brownlee

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

The complete structure of the neuraminidase gene in influenza A/PR/8/34 has been determined by cloning into the bacteriophage M13 and sequencing with dideoxynucleotide chain terminators. The gene is 1,413 nucleotides long, codes for a protein of 454 amino acids and has five potential glycosylation sites. We suggest that the neuraminidase, unlike the influenza haemagglutinin, is oriented with its N-terminus buried in the viral membrane.

INFLUENZA A viruses have two surface antigens, haemagglutinin and neuraminidase, both of which can undergo extensive antigenic variation. The haemagglutinin binds the virus to sialic acid-containing receptors on the cell, but the role of the neuraminidase is still not fully understood (for reviews see refs 1, 2). The enzyme cleaves terminal *N*-acetylneuraminic acid (sialic acid) from carbohydrate chains in glycoproteins³, and its function may be to remove the carbohydrate residues from the virus envelope itself, thus eliminating receptors for the haemagglutinin and preventing self-aggregation of virion particles⁴. The removal of sialic acid also seems to facilitate proteolytic cleavage of the haemagglutinin, which is essential for infectivity^{5,6}, by exposing the site of cleavage on the haemagglutinin^{7,8}.

Although the haemagglutinin is the more significant factor in the immune response of the host organism, the neuraminidase is also involved in this response^{9,10}, as anti-neuraminidase antibodies can inhibit viral replication. Like the haemagglutinin, the neuraminidase can undergo either small changes in antigenicity^{11,12} (antigenic drift) or major changes (antigenic shift), in which the neuraminidases do not cross-react serologically¹¹. Antigenic shift in human influenza viruses occurred in 1957, when the N2 subtype replaced the N1. A comparison of the N-terminal amino acid sequences predicted from RNA sequencing¹³ showed that within a subtype there is a high degree of amino acid conservation, but between the different subtypes there is little homology after the first 12 amino acids.

Using recombinant DNA methods, we have sequenced the neuraminidase of the human strain A/PR/8/34 (H1N1, formerly designated H0N1) to provide a basic framework for both understanding the function of the protein and to further study the mechanism of antigenic drift and shift. From the sequence has emerged the possibility that in contrast to the haemagglutinin^{14–20}, the neuraminidase is bound to the viral membrane by its N-terminus.

Cloning and sequence analysis in M13

Three separate strategies were used to determine the sequence. As part of a 'shotgun' approach²¹ to sequence the eight single-stranded RNA segments in the influenza genome^{22–24}, restriction fragments derived from the whole genome were cloned into

the bacteriophage M13mp2 (ref. 25) or one of its derivatives^{26,27} and sequenced by the dideoxy technique²⁸. One *Hae*III clone contained the sequence of 444 nucleotides at the 3' end of the neuraminidase virion RNA (vRNA) as seen from a match with limited RNA sequence data for the related fowl plague virus (FPV) neuraminidase gene^{29,30}. Two other clones, derived from

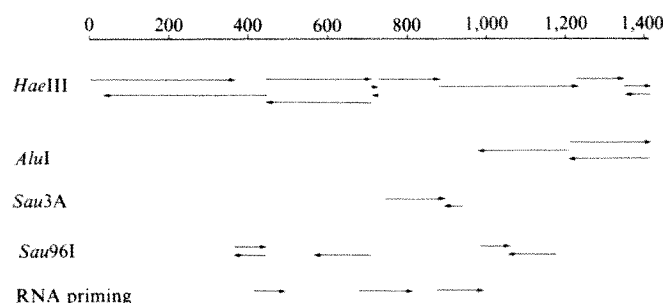


Fig. 1 Summary of sequence evidence. Virion RNA was polyadenylated and double-stranded cDNA prepared using oligo(dT)_{12–18} priming and then loopback of the 3' end of the cDNA as described elsewhere²¹. Alternatively, the band 6 cDNA derived from 20 µg RNA was eluted from a 2.8% acrylamide/7 M urea gel³⁴ and annealed in 40 µl of 0.05 M NaCl, 0.01 M Tris-HCl pH 7.4, 0.01 M MgCl₂, 1 mM dithiothreitol (DTT) to 80 pmol of a 13-long synthetic oligodeoxynucleotides primer (see text) by boiling and then slowly cooling. Deoxynucleotide triphosphates to 0.5 mM and 2U DNA polymerase (Klenow fragment, Boehringer) were added and the reaction incubated for 1 h at 20 °C. Double-stranded molecules were cut by restriction enzymes; *Hae*III and *Alu*I fragments were blunt-end ligated to 10 ng *Hind*III-cut M13mp7 (ref. 27) in 10 µl of 0.05 M Tris-HCl pH 7.4, 0.01 M MgCl₂, 0.01 M DTT, 1mM rATP and 0.4U T4 DNA ligase. *Sau*96I fragments were first blunt-ended by incubation with the Klenow fragment of DNA polymerase³⁶ and then ligated as above. *Sau*3A fragments were ligated as described elsewhere²¹. DNA sequencing of M13 clones used a 30-base pair cloned primer⁶⁵ or 17-long oligodeoxynucleotide synthetic primer (from Dr M. J. Gait). RNA priming with restriction fragment primers was as described elsewhere²¹. Sequence reactions were electrophoresed on 6% acrylamide/7 M urea thin gels⁶⁶.

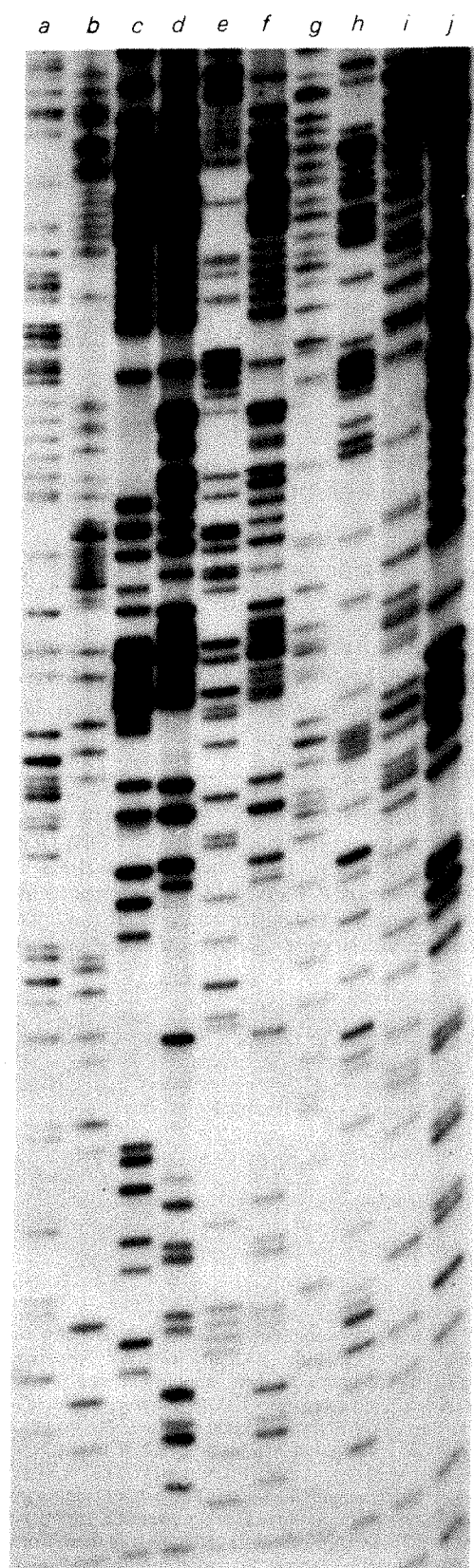


Fig. 2 Screening clones by dideoxythymidine tracks. *Hae*III clones were isolated as described in Fig. 1 legend. Screening using dideoxythymidine patterns as characterization³⁶ was on a 6% acrylamide/7 M urea thin gel⁶⁶ electrophoresed at 30 mA for 1.5 h. The 10 clones screened on this gel correspond to the following sequences (see Fig. 3): *a*, positions 1,347 to 1,413; *b*, 1,346–1,233; *c*, 1,346–1,233; *d*, 1,233–1,346; *e*, 727–878; *f*, 1,233–1,346; *g*, 710–726; *h*, 879–1,232; *i*, 710–726; and *j*, 1,347–1,413. Therefore clones *a* and *j* are identical, as are *g* and *i*. Clones *b* and *c* contain the same *Hae*III fragment but the opposite strand to clones *d* and *f*.

Table 1 Amino acid composition of the A/PR/8/34 neuraminidase

Amino acid	No.	Amino acid	No.
Ala (A)	16	Leu (L)	21
Arg (R)	20	Lys (K)	23
Asn (N)	25	Met (M)	7
Asp (D)	25	Phe (F)	16
Cys (C)	19	Pro (P)	21
Glu (E)	17	Ser (S)	51
Gln (Q)	11	Thr (T)	29
Gly (G)	44	Trp (W)	16
His (H)	10	Tyr (Y)	14
Ile (I)	41	Val (V)	28
Total no. amino acids 454			

*Hae*III and *Alu*I shotgun digests, contained an additional 500 nucleotides of neuraminidase sequence when overlapped by the other approaches.

To extend the 3' sequence, we prepared double-stranded cDNA (dsDNA) specific for segment 6, the neuraminidase gene^{31–33}, after first purifying the single-stranded cDNA (ssDNA) on an acrylamide gel³⁴. Clones derived from restriction digests of this dsDNA (Fig. 1) provided all of the remaining sequence of the neuraminidase gene, but not every overlap. The *in vitro* back-copy reaction was greatly facilitated by the use of a 13-long oligodeoxynucleotide primer³⁵ complementary to the 3' termini of the ssDNAs²⁹, which allowed synthesis of the complete sequence corresponding to the 5' end of the vRNA.

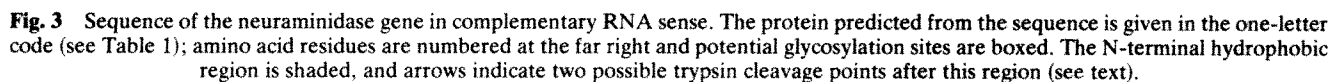
The final three overlaps were obtained from the third approach, in which restriction fragments derived from the replicative form of selected recombinant M13 clones were used as primers for dideoxy sequencing on a total vRNA template²¹.

Our approach was such that most sequences were determined more than once. As clones were obtained in either orientation in M13, 93% of the neuraminidase gene was determined on both strands, and of the remainder, most was sequenced from at least two independent clones. A total of 25 clones and three RNA primings was used in the complete sequence determination.

Because no clone is identified before its sequencing, particular restriction digests may yield an individual fragment in the same orientation in many clones. This problem of redundant information, inherent in a 'shotgun' strategy, was avoided by eliminating excess clones with a rapid characterization³⁶ that used only one of the four dideoxy tracks of a normal sequence reaction. Figure 2 shows the pattern of 'dideoxythymidine tracks' for one set of *Hae*III fragments derived from band 6 dsDNA. Such tracks show whether a clone has been obtained before and can indicate the length of an insert by inspection of the pattern for the M13 bands which follow the insert.

Nucleotide sequence of the neuraminidase gene

Figure 3 shows the complete sequence of the neuraminidase gene as the complementary RNA, corresponding to the messenger sense. The sequence is 1,413 nucleotides long, including 20 nucleotides 5' to the first AUG and 31 nucleotides 3' to the last codon. (The actual structure of the mRNA produced by the virus has the cap and several nucleotides donated by a host mRNA³⁷ and is an incomplete transcript of the vRNA³⁸.) Only one reading frame is open, beginning with the AUG at position 21 and terminating at a UAG at position 1,383; the next in-frame AUG is at position 450 and could code for a protein with a molecular weight (MW) of only 34,000 (see below). The other two reading frames are blocked by numerous termination codons. The predicted protein is co-linear with the RNA and occupies 97% of the coding capacity of the gene. Whereas band 8 (refs 39, 40) and possibly band 77 (ref. 21) of influenza virus code for two proteins, the neuraminidase gene is similar to the haemagglutinin gene^{16–20} in that it codes for a single protein. Previous sequencing of the A/PR/8/34 neuraminidase gene has determined limited sequences at the 3'



The influenza virus neuraminidase is an integral membrane protein and can only be released from the viral envelope by detergent or protease treatment⁵⁰. Previously it was suggested⁴⁵



Fig. 4 Comparison of membrane attachment sites. The N-terminal hydrophobic sequence of the neuraminidase is compared with the membrane attachment regions of isomaltase⁵⁹, glycophorin⁵⁴, vesicular stomatitis virus (VSV) glycoprotein⁵⁷, membrane-bound IgM⁵⁶ and the influenza virus haemagglutinin of A/PR/8/34 (unpublished data). The boxed region indicates the uncharged residues thought to span the membrane. The sequences have been aligned to show amino acid homology with the neuraminidase.

that a hydrophobic region was present at either the N- or C-terminus because the trypsin-released enzyme is smaller (by an estimated MW of 7,000 (ref. 45) or 12,000 (ref. 43)) than that derived from SDS disruption of the virion. This hydrophobic region would anchor the neuraminidase to the lipid bilayer of the virus. An examination of the A/PR/8/34 neuraminidase (Fig. 3) shows only one major hydrophobic region, located near the N-terminus and extending from amino acid residue 7 to 35. Of these 29 residues, 18 are hydrophobic, 11 are neutral and none is charged. The first trypsin-sensitive residues following this hydrophobic region are at positions 57(Lys) and 63 (Lys), and tryptic cleavage at either of these positions would result in a loss of MW of ~6,000, which is largely consistent with protein chemistry studies^{43,45}. A stretch of 16 uncharged amino acid residues (420–435) lies near the C-terminus, but of these only 5 are hydrophobic and seem unlikely to be sufficient for attachment of an integral membrane protein. Additional evidence favouring an N-terminal attachment comes from the finding that the small fragment lost on protease isolation of the enzyme from B/Lee⁴⁵ or FPV⁵¹ is rich in carbohydrate. In the A/PR/8/34 neuraminidase, potential carbohydrate attachment sites near the N-terminus occur at residues 44 and 58, and at least the first of these sites would be contained within the tryptic peptide presumed to be lost. No attachment sites lie within the C-terminal half of the molecule.

This membrane orientation would contrast with that of the influenza virus haemagglutinin⁵², and also of such membrane glycoproteins as erythrocyte glycophorin^{53,54}, HLA antigens⁵⁵, membrane-bound immunoglobulin μ chains⁵⁶ and vesicular stomatitis virus glycoprotein⁵⁷, all of which have C-terminal hydrophobic sequences. However, attachment of proteins to membranes through a hydrophobic N-terminus does have a precedent in the orientation of intestinal brush border aminopeptidase⁵⁸ and isomaltase^{58–60}. Figure 4 shows a comparison of the N-terminal hydrophobic sequence of the neuraminidase with other sequences involved in spanning the lipid bilayer; glycophorin shows the greatest sequence homology to the neuraminidase but all the sequences are similarly hydrophobic.

There is no experimental evidence to indicate whether the N-terminal hydrophobic sequence of the neuraminidase is a 'signal sequence' which is cleaved off⁶¹, or whether it remains

attached to the membrane, as has been suggested for the N-terminal region of isomaltase⁵⁹. However, in the two human subtypes, the total conservation of the first 12 amino acid residues¹³, of which the first six are hydrophilic, suggests that these have some critical role, which might be to interact with another protein or neuraminidase subunit. (Signal sequences of the haemagglutinins, which are cleaved off⁶², show hydrophobicity²⁰ but no conservation of amino acid sequence.) We suggest that the neuraminidase, like isomaltase, has an 'extended signal sequence'⁵⁹ which both transfers the protein across the membrane and remains in the bilayer to anchor the protein, with the first six residues protruding into the virion as a hydrophilic tail. The signal sequence may not be cleaved because the recognition sequence for the signal peptidase is lacking or unexposed⁵⁹. Such an N-terminal attachment favours a loop model⁶³ of protein translocation across a membrane, in which the signal sequence inserts into the lipid bilayer as a loop, with the first few amino acid residues remaining on the cytoplasmic side of the membrane. This hypothesis concerning N-terminal attachment of the neuraminidase could be tested directly by amino acid sequence analysis of the intact and trypsin-released enzyme.

Concluding remarks

The sequence of the neuraminidase gene allows us to suggest a possible role for a long hydrophobic sequence at the protein's N-terminus, but the regions involved in enzymatic activity and antigenicity remain undefined. Future sequences of other neuraminidase genes should reveal regions of homology within and between subtypes, and thus elucidate the phenomena of antigenic drift and shift at the molecular level. A more complete understanding of the neuraminidase, however, must await a three-dimensional structure of the molecule, which is at the preliminary stage for the N2 subtype⁶⁴.

We thank Mrs M. A. Robertson for supplying virus and RNA, R. Henderson, M. Bretscher and D. L. Bentley for discussions, A. D. McLachlan for the Chou–Fasman prediction and J. Messing for M13 vectors. S.F. was supported by a NSF predoctoral fellowship and G.W. by a research fellowship from Trinity College, Cambridge.

Received 7 November 1980; accepted 15 January 1981.

- Schulze, I. T. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 53–82 (Academic, London, 1975).
- Bucher, D. & Palese, P. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 83–123 (Academic, London, 1975).
- Gottschalk, A. *Biochim. biophys. Acta* **23**, 645–646 (1957).
- Palese, P. *et al. Virology* **61**, 397–410 (1974).
- Klenk, H.-D. *et al. Virology* **68**, 426–439 (1975).
- Lazarowitz, S. G. & Choppin, P. W. *Virology* **68**, 440–454 (1975).
- Schulman, J. L. & Palese, P. *J. Virol.* **24**, 170–176 (1977).
- Nakajima, S. & Sugiura, A. *Virology* **101**, 450–457 (1980).
- Schulman, J. L., Khakpour, M. & Kilbourne, E. D. *J. Virol.* **2**, 778–786 (1968).
- Rott, R., Becht, H. & Orlich, M. *J. gen. Virol.* **22**, 35–41 (1974).

- Paniker, C. K. *J. gen. Virol.* **2**, 385–394 (1968).
- Schulman, J. L. & Kilbourne, E. D. *Proc. natn. Acad. Sci. U.S.A.* **63**, 326–333 (1969).
- Blok, J. & Air, G. M. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 125–133 (Elsevier, Amsterdam, 1980).
- Ward, C. W. & Doppeide, T. A. *Br. med. Bull.* **35**, 51–56 (1979).
- Waterfield, M. D. *et al. Br. med. Bull.* **35**, 57–63 (1979).
- Porter, A. G. *et al. Nature* **282**, 471–477 (1979).
- Min Jou, W. *et al. Cell* **19**, 683–696 (1980).
- Both, G. W. & Sleight, M. J. *Nucleic Acids Res.* **8**, 2561–2575 (1980).
- Verhoeven, M. *et al. Nature* **286**, 771–776 (1980).
- Gething, M.-J. *et al. Nature* **287**, 301–306 (1980).
- Winter, G. & Fields, S. *Nucleic Acids Res.* **8**, 1965–1974 (1980).
- McGeoch, D. J., Fellner, P. & Newton, C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3045–3049 (1976).

23. Palese, P. & Schulman, J. L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2142-2146 (1976).
24. Scholtissek, C. *et al. Virology* **74**, 332-344 (1976).
25. Gronenborn, B. & Messing, J. *Nature* **272**, 375-377 (1978).
26. Rothstein, R. J. *et al. Meth. Enzym.* **68**, 101-110 (1979).
27. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309-321 (1981).
28. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
29. Skehel, J. J. & Hay, A. J. *Nucleic Acids Res.* **5**, 1207-1219 (1978).
30. Robertson, J. S. *Nucleic Acids Res.* **6**, 3745-3757 (1979).
31. Palese, P. & Schulman, J. L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2142-2146 (1976).
32. Scholtissek, C. *et al. Virology* **74**, 332-344 (1976).
33. Inglis, S. C., McGeoch, D. J. & Mahy, B. W. J. *Virology* **78**, 522-536 (1977).
34. Sleight, M. J., Both, G. W. & Brownlee, G. G. *Nucleic Acids Res.* **6**, 1309-1321 (1979).
35. Winter, G., Fields, S., Gait, M. J. & Brownlee, G. G. *Nucleic Acids Res.* **9**, 237-245 (1981).
36. Sanger, F. *et al. J. molec. Biol.* **143**, 161-178 (1980).
37. Krug, R. M., Broni, B. A. & Bouloy, M. *Cell* **18**, 329-334 (1979).
38. Hay, A. J. *et al. Nucleic Acids Res.* **4**, 4197-4209 (1977).
39. Inglis, S. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 3790-3794 (1979).
40. Lamb, R. A. & Choppin, P. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4908-4912 (1979).
41. Desselberger, U., Zamecnik, P. & Palese, P. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 169-179 (Elsevier, Amsterdam, 1980).
42. Desselberger, U. *et al. Gene* **8**, 315-328 (1980).
43. Wrigley, N. G. *et al. Virology* **51**, 525-529 (1973).
44. Gregoriades, A. *Virology* **49**, 333-336 (1972).
45. Lazdins, I., Haslam, E. A. & White, D. O. *Virology* **49**, 758-765 (1972).
46. Laver, W. G. & Baker, N. J. *gen. Virol.* **17**, 61-67 (1972).
47. Neuberger, A. *et al. in The Glycoproteins: Their Composition, Structure and Function* (ed. Gottschalk, A.) 450-490 (Elsevier, Amsterdam, 1972).
48. Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 211-245 (1974).
49. Schiffer, M. & Edmundson, A. B. *Biophys. J.* **7**, 121-135 (1967).
50. Drzeniek, R., Seto, J. & Rott, R. *Biochim. biophys. Acta* **128**, 547-558 (1966).
51. Rott, R. *et al. Naturforsch. B.* **27**, 227-233 (1972).
52. Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 93-97 (1975).
53. Tomita, M. & Marchesi, V. T. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2964-2968 (1975).
54. Furthmayr, H. *et al. Archs Biochem. Biophys.* **185**, 21-29 (1978).
55. Robb, R. J., Terhorst, C. & Strominger, J. L. *J. biol. Chem.* **253**, 5319-5324 (1978).
56. Rogers, J. *et al. Cell* **20**, 303-312 (1980).
57. Rose, J. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3884-3888 (1980).
58. Maroux, S. & Louvard, D. *Biochim. biophys. Acta* **419**, 189-195 (1976).
59. Frank, G. *et al. FEBS Lett.* **96**, 183-188 (1978).
60. Brunner, J. *et al. J. biol. Chem.* **254**, 1821-1828 (1979).
61. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835-851 (1975).
62. Elder, K. T. *et al. Virology* **95**, 343-350 (1979).
63. Inouye, M. & Halegoua, S. *CRC crit. Rev. Biochem.* **7**, 339-371 (1980).
64. Colman, P. M., Tulloch, P. A. & Laver, W. G. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 351-356 (Elsevier, Amsterdam, 1980).
65. Anderson, S. *et al. Nucleic Acids Res.* **8**, 1731-1743 (1980).
66. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107-110 (1978).

A plasmid DNA primase active in discontinuous bacterial DNA replication

Brian M. Wilkins & Graham J. Boulnois

Department of Genetics, University of Leicester, Leicester LE1 7RH, UK

Erich Lanka

Max-Planck-Institut für Molekulare Genetik, Abteilung Schuster, Berlin-Dahlem, FRG

A DNA primase encoded by an IncIα plasmid promotes efficient DNA replication in a primase-defective mutant of Escherichia coli. This finding implies that the plasmid enzyme can prime discontinuous DNA synthesis of the bacterial chromosome. The plasmid gene encodes two large, antigenically related proteins which differ from E. coli primase.

A CENTRAL feature of DNA replication is the discontinuous synthesis of the strand that grows in the overall direction of 3' to 5'. DNA polymerases cannot initiate synthesis of the fragments involved in the process because these enzymes require the 3'-OH terminus of an oligonucleotide primer (for review see ref. 1). During replication of the *Escherichia coli* chromosome, the primer sequences are synthesized by the product of the *dnaG* gene², called primase³. This 60,000 molecular weight (*M_r*) single subunit protein⁴ is thought to act in a multi-enzyme complex that includes *dnaB* product⁵. Although certain virulent bacteriophages are known to encode DNA primases, these enzymes act in phage DNA replication in cooperation with specific phage DNA polymerases⁶⁻⁹.

Conjugative plasmids of the IncIα and IncIγ incompatibility groups¹⁰, represented by Coll, R64, R144 and R621a, can suppress temperature-sensitive *dnaG* mutations, allowing DNA replication and colony formation at high temperature^{11,12}. The relevant gene is designated¹³ *sog*, conferring a phenotype of Sog⁺. Suppression of *dnaG* lesions requires derepression of the plasmid transfer genes by a *drd* mutation, implying that the physiological role of *sog* product may be in conjugation. Recently, a novel DNA primase has been purified from bacteria containing the IncIα plasmid R64*drd*-11 and characterized¹⁴. The enzyme has a low template specificity *in vitro*; it can substitute for the activities of *dnaB*-*dnaC*-*dnaG* products, *E. coli* RNA polymerase and *dnaG* product in priming synthesis of DNA complementary to viral strands of phages ΦX174, fd and G4, respectively.

The studies reported here show that the IncIα plasmid product responsible for suppressing bacterial *dnaG* mutations is a DNA primase. This finding implies that the plasmid enzyme can substitute for *E. coli* primase to promote efficient discontinuous synthesis of the bacterial chromosome and indicates that primer-generating enzymes other than *dnaG* protein can act in bacterial DNA replication. The approach has involved the

cloning of the *sog* gene of Coll*drd*-1 and analysis of proteins encoded by the parental and recombinant plasmids. The *sog* gene is shown to encode two antigenically related proteins. Primase activity is localized in part of the larger molecule.

Table 1 Plasmids that suppress *dnaG3* specify the IncIα plasmid primase

Strain	Colony formation 40 °C/30 °C	DNA synthesis at 41 °C (10 ⁻³ c.p.m. per ml per A ₄₅₀ unit)	Primase activity (mU per mg CCE)
BW86	<10 ⁻⁷	11	ND
BW86 (pBR325)	<10 ⁻⁷	16	ND
BW86 (Coll)	2.7 × 10 ⁻⁷	22	1
BW86 (Coll <i>drd</i> -1)	2.1 × 10 ⁻¹	342	213
BW86 (pLG215)	1.8 × 10 ⁻⁴	262	182
BW86 (pLG214)	1.0	341	221

Colony formation gives the fractional yield of colonies at the indicated temperatures. Overnight cultures in nutrient broth were diluted suitably and plated on prewarmed nutrient agar. Thymine (20 μg ml⁻¹) was added to all media. DNA synthesis indicates ³H-thymine incorporated in 60 min at 41 °C. Bacteria, growing exponentially in SGC medium containing deoxyguanosine and thymine (3 μg ml⁻¹)¹¹, were incubated at 41 °C for 5 min before addition of ³H-thymine to 7.5 μCi ml⁻¹ (320 mCi mmol⁻¹). A₄₅₀ of cultures was measured at zero time. Primase was assayed using bacteria growing exponentially in broth. 2 × 10⁹ bacteria were gently lysed using the non-ionic detergent Brij 58 in conjunction with lysozyme²⁶. Following centrifugation (80,000 g, 30 min), primase activity and protein concentration were measured as described previously¹⁴. All liquid media for strains harbouring pBR325, pLG214 or pLG215 contained tetracycline (10 μg ml⁻¹) and agar contained ampicillin (25 μg ml⁻¹). 1 unit of primase incorporated 1 μmol of dTMP in the conditions used. CCE, crude cell extract; ND, not detectable.

Cloning of the *sog* gene

The small multicopy plasmid pBR325 was used as the vector for cloning. This plasmid¹⁵ carries genes conferring resistance to ampicillin (Ap^r), tetracycline (Tc^r) and chloramphenicol (Cm^r) and contains a single endonuclease *EcoRI* cleavage site in the latter. Fragments of *ColIdrd-1*, generated by *EcoRI* restriction endonuclease, were ligated into this site in pBR325 and the DNA was used to transform a temperature-sensitive *dnaG3* mutant (BW86, *dnaG3 leu thyA deoB rpsL ΔchlA-uvrB ColI^r*; ref. 13). Transformants harbouring *Sog*⁺ plasmids could be isolated using the following strategy. Ap^r transformants were first selected at 30 °C and allowed to grow into small colonies, when they were replica plated. About 1% of the replicas grew at 40 °C to give distinct colonies; these were found to contain *Sog*⁺ recombinant plasmids.

The plasmid (pLG211) in one of these strains was characterized. Cleavage with *EcoRI* restriction endonuclease generated three DNA fragments with estimated sizes of 5.9 kilobases (which corresponds to the vector), 7.9 and 0.27 kilobases. The smallest of these DNA fragments was next removed by cleaving pLG211 with *EcoRI* endonuclease followed by ligation to give pLG215. The latter plasmid specifies resistance to ampicillin

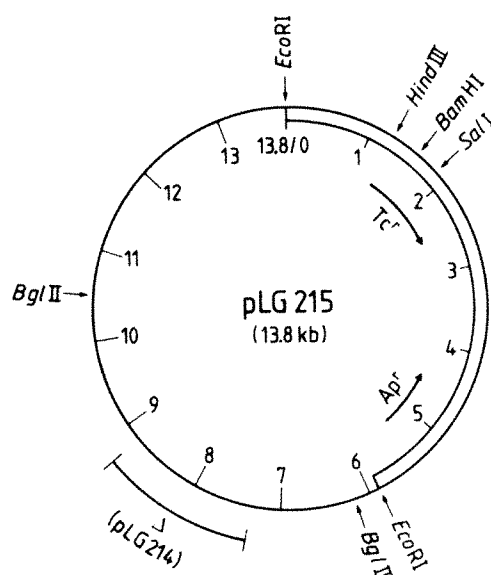


Fig. 1 Map of pLG215 showing the position of the 1.5-kilobase deletion mutation present in pLG214. The vector, pBR325, is drawn in double lines (0 to 5.9 kilobases). The positions of the Tc^r and Ap^r genes are derived from maps of pBR322 and pBR325 (refs 15, 27). Mapping was carried out by subjecting restriction endonuclease-cleaved DNA to electrophoresis in 0.5% or 2% agarose gels using procedures similar to those described elsewhere^{28,29}, except that double digestions were carried out in 6.6 mM $MgCl_2$, 60 mM NaCl, 6.6 mM 2-mercaptoethanol and 6.6 mM Tris-HCl (pH 7.4). Molecular weight markers included phage λ DNA cleaved with *EcoRI* and *HindIII* endonucleases in single or double digests³⁰ and *Taq*YI-cleaved pBR322 (ref. 31). Lengths (kilobases) of DNA fragments in *Bgl*II+*EcoRI* double digests were 5.9, 4.4, 3.3 and 0.23. Lengths of DNA fragments generated by *Bgl*II+*HindIII* endonucleases were 4.9, 4.6 and 4.4 kilobases. The 4.4-kilobase DNA fragment was replaced by a 2.9-kilobase fragment in corresponding digests of pLG214. The *Bam*HI and *Sal*I cleavage sites were mapped using sizes of fragments generated in appropriate double digests with *HindIII* or *EcoRI* endonucleases. Construction of pLG215 is summarized in the text. pBR325 was isolated by a cleared lysate method³². *ColIdrd-1* was purified using a Sarkosyl lysis method³³ with the modification that DNA in supernatant was concentrated by ethanol precipitation following extraction with phenol. Ligations were done at 4 °C in 60 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.1 mM ATP, 0.005% gelatin and T4 DNA ligase (from Dr Ron Wilson).

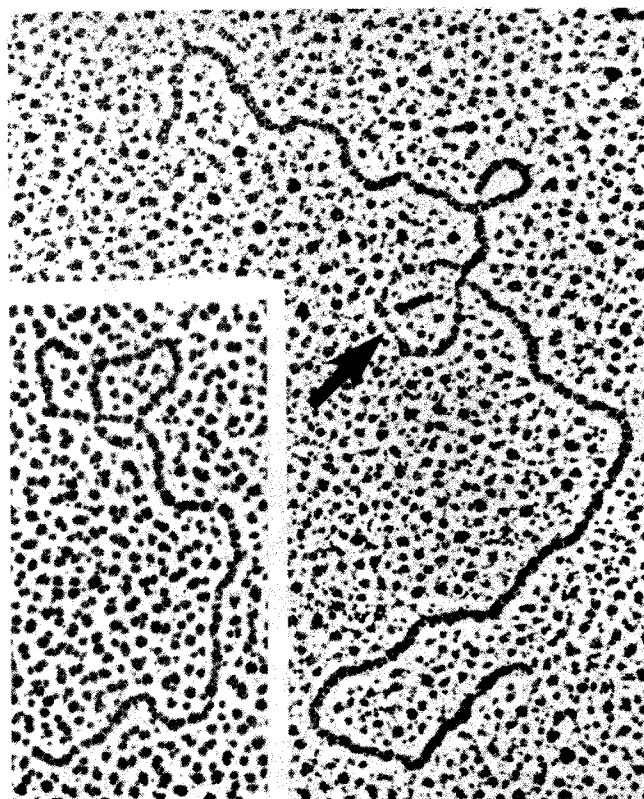


Fig. 2 Electron micrograph of a pLG214/pLG215 heteroduplex DNA molecule showing a deletion loop, indicated by the arrow. The plasmids were cleaved at the unique *HindIII* site (Fig. 1) before formation of heteroduplexes. The mean fractional length of the short arm of 12 molecules was 0.48. The inset shows a heteroduplex DNA molecule prepared following digestion of pLG214 and pLG215 with *EcoRI* endonuclease. This digestion separates the cloned fragment of *ColIdrd-1* from the vector (Fig. 1). The mean fractional length of the short arm of 18 molecules was 0.22 (s.d. $\pm$ 3.7%) of the duplex DNA. The estimated length of the duplex DNA is 6.4 kilobases, which is the size of the relevant *EcoRI* fragment in pLG214. Molecules were prepared and examined using methods described elsewhere³⁴.

and to tetracycline but not to chloramphenicol, indicating insertion-inactivation of the Cm^r gene of pBR325. Plasmid pLG215 retains the 7.9-kilobase *EcoRI* restriction fragment of *ColIdrd-1* which was found to contain two endonuclease *Bgl*II cleavage sites (Fig. 1). Table 1 shows that the plasmid suppresses the *dnaG* mutation in BW86 as measured by extensive DNA synthesis at 41 °C and colony formation at 40 °C. However, by both criteria, pLG215 suppressed the *dnaG* lesion less efficiently than *ColIdrd-1*.

BW86(pLG215) is deficient in accumulating DNA at high temperature. This was established by monitoring absorbance and trichloroacetic acid (TCA)-precipitable ³H-thymine in a uniformly labelled culture: following transfer of an exponentially growing culture from 31 °C to 41 °C, A_{450} increased by a factor of 7.0 in 120 min at high temperature whereas DNA increased by a factor of 3.6. No such disparity was detected at 31 °C. This finding, together with the heterogeneous size of colonies formed at 40 °C, suggested that an undefined mutation might be necessary to allow BW86 (pLG215) to form colonies at high temperature. Chromosomal mutation is inferred from studies with six independently isolated colonies formed at 40 °C. The plasmids in these strains appeared identical to pLG215 as judged by analysis of *EcoRI* restriction digests and, following reintroduction of the plasmids into BW86, by the ability of the resulting transformants to form colonies at 40 °C. Plasmid-free segregants of the six original strains were temperature sensitive but they acquired a fully thermoresistant phenotype with high frequency in a second round of transformation with pLG215.

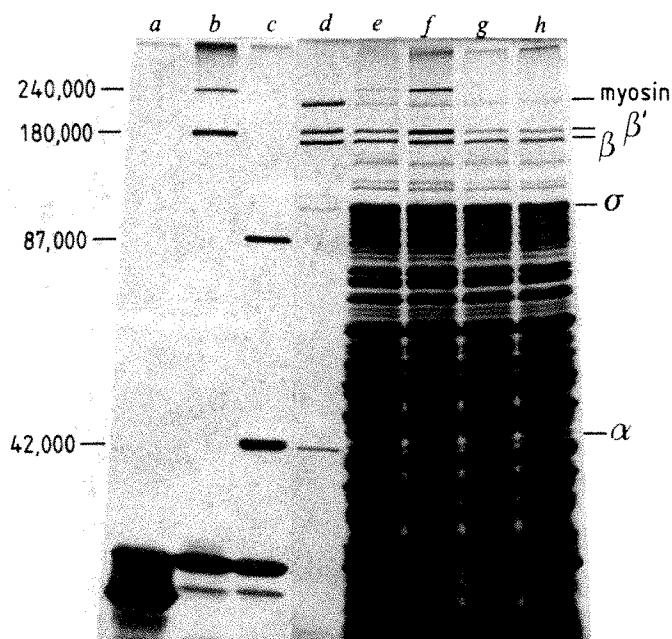


Fig. 3 Analysis by SDS-gel electrophoresis of proteins encoded by *Sog*⁺ plasmids. Lanes *a*–*c* show a fluorograph of ³⁵S-labelled proteins synthesized in minicells containing: *a*, pBR325; *b*, pLG215; *c*, pLG214. Lanes *e*–*h* show Coomassie blue-stained proteins in extracts of: *e*, BW86 (pLG215); *f*, BW84 (pLG216); *g*, BW86 (pLG214); *h*, BW86 (pBR325). Molecular weight standards (lane *d*) are myosin (*M_r* 212,000), *E. coli* RNA polymerase β'-subunit (*M_r* 165,000), β-subunit (*M_r* 155,000), σ-subunit (*M_r* 90,000) and α-subunit (*M_r* 39,000). Minicells were purified³⁵ from plasmid-containing derivatives of DS410 (ref. 36). Proteins were labelled for 90 min at 37 °C in M9 minimal medium containing ³⁵S-methionine (100 μCi ml⁻¹, 1 Ci mmol⁻¹) and Difco methionine assay medium (1.6 mg ml⁻¹). After a 10-min chase with methionine-supplemented broth, minicells were lysed by boiling in cracking buffer (62.5 mM Tris-HCl (pH 6.8), 3% SDS, 10% glycerol, 5% β-mercaptoethanol). About 150,000 c.p.m. were added per slot. Conditions for gel electrophoresis³⁷ and fluorography³⁸ have been described. Cell extracts were prepared from 1 ml of stationary phase bacteria at *A*₄₅₀ of 1. Cells were lysed in 30 μl cracking buffer and the total lysate was added to the slot.

Thus, the strains contain a chromosomal mutation which does not revert the *dnaG3* lesion but allows efficient colony formation at 40 °C in the presence of the *Sog*⁺ plasmid.

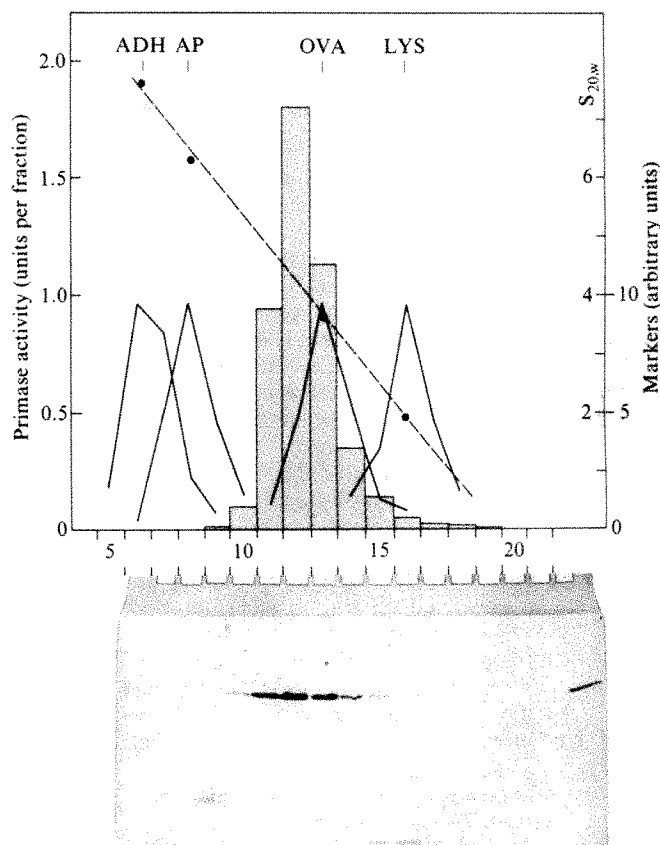
The recloning experiment which generated pLG215 from pLG211 gave rise to a BW86 transformant strain that was atypical in forming colonies at 40 °C with a frequency of 1. This phenotype was determined by a plasmid designated pLG214 (Table 1). Endonuclease *Eco*RI cleaved pLG214 to produce DNA fragments of 5.9 and 6.4 kilobases, suggesting deletion of 1.5 kilobases from the segment of *Colldrd-1* present in pLG211 and pLG215. Endonuclease *Bgl*II digests showed that the deletion affects the 4.4-kilobase *Bgl*II fragment found in pLG215 (Fig. 1). This location was confirmed by analysis of

pLG214/pLG215 heteroduplex DNA molecules which demonstrated a deletion loop almost equidistant from the termini generated by cleavage of the plasmids at the unique *Hind*III site (Fig. 2). Heteroduplex DNA formed between endonuclease *Eco*RI-cleaved pLG214 and pLG215 indicated a single internal deletion mapping about 1.4 kilobases from the nearer junction of the cloned DNA and vector (Fig. 2 inset).

Sog⁺ recombinant plasmids specify DNA primase

Table 1 shows that bacteria harbouring *Sog*⁺ recombinant plasmids produce DNA primase. The assay measured the ability

Fig. 4 Glycerol gradient centrifugation of DNA primase specified by pLG214. 47 μg of protein in 0.2 ml of gradient buffer was layered on a 3.7 ml, 18–43% (wt/vol) glycerol gradient in 20 mM Tris-HCl (pH 7.6), 200 mM NaCl, 1 mM dithiothreitol and 1 mM EDTA. Centrifugation was for 24 h in the Beckman SW-60 rotor at 60,000 r.p.m. Sedimentation was from right to left. 173-μl fractions were assayed for primase activity (shaded histogram). Aliquots (150 μl) were subjected to SDS-gel electrophoresis (lower panel). The lane at the right contained 1.9 μg of primase preparation. Yeast alcohol dehydrogenase (ADH, *M_r* 141,000), *E. coli* alkaline phosphatase (AP, *M_r* 80,000), ovalbumin (OVA, *M_r* 43,000) and lysozyme (LYS, *M_r* 13,970) were run in parallel as markers. Primase was purified from BW86 (pLG214). 280 g of wet cell paste in 50 mM Tris-HCl (pH 7.6), 100 mM NaCl and 5% sucrose was frozen and thawed and then supplemented with EDTA (1 mM), spermidine-3HCl (5 mM), Brij 58 (0.25%) and lysozyme (0.04%) at a final concentration of 1 g cells per 2.5 ml. After 90 min at 0 °C the suspension was centrifuged (90 min, 70,000 g). All subsequent buffers were supplemented with 10% glycerol, 1 mM EDTA and 1 mM dithiothreitol. Supernatant containing 8.3 g of protein was added to a heparin-Sepharose CL6B column and eluted by applying a 600 ml gradient of 0.05–1 M NaCl in 20 mM Tris-HCl (pH 7.6). Primase activity eluted at 0.5 M NaCl. Active fractions (containing 86 mg of protein) were applied to a hydroxyapatite column and protein was eluted by a 300-ml gradient of potassium phosphate pH 7.0 (20–400 mM). Enzyme activity eluted at 150 mM phosphate. Active fractions (42.5 mg of protein) were applied to a phosphocellulose column (Whatman P11) and a 300-ml 0–0.8 M NaCl gradient was applied which eluted primase at 300 mM. After dialysis against buffer A (20 mM Tris-HCl pH 7.6, 50 mM NaCl), 30 ml (8.1 mg of protein) was applied to a DEAE-Sephacel column. Primase activity failed to bind and was washed off with buffer. This fraction (3.9 mg of protein) was applied to a second phosphocellulose column. Active fractions were concentrated by dialysis against 20% polyethylene glycol in buffer A. Yield was 2.6 mg of protein and 80% of activity. The enzyme remains stable at –20 °C for at least 6 months.



of crude cell extracts to stimulate conversion of single-stranded DNA of phage fd to duplex replicative form in a receptor extract containing rifampicin¹⁴. The antibiotic was added to inhibit RNA polymerase, the enzyme naturally required for priming DNA synthesis on viral strands of phage fd¹⁶. A 200-fold higher activity was detected in ColI-containing strains when the plasmid carried the *drd* mutation. This correlates with the requirement for a *drd* mutation in plasmid-mediated suppression of the *dnaG* lesion. *Sog*⁺ recombinant plasmids pLG214 and pLG215 directed synthesis of similar high activities of the primase, indicating that the gene specifying the enzyme is in the same *EcoRI* fragment of ColI*drd-1* as the gene that causes suppression of *dnaG* mutations.

Polypeptides encoded by *Sog*⁺ plasmids

Plasmids pLG214 and pLG215 segregate into chromosomeless minicells, providing a system for identification of polypeptides encoded by the cloned fragment of ColI*drd-1* that contains the *sog* gene. Comparison of ³⁵S-methionine-labelled proteins synthesized in pBR325- and pLG215-containing minicells (Fig. 3, lanes a, b) shows that the 7.9-kilobase *EcoRI* restriction fragment of ColI*drd-1* encodes two polypeptides of *M*_r 240,000 and 180,000. Assuming these to be encoded totally by the ColI component of the recombinant plasmid, coding sequences comprising about 0.82 and 0.62 of the DNA, respectively, will be required. The sum of these values implies that the genetic determinants of the two polypeptides share common nucleotides. Lane c indicates that the 6.4-kilobase deletant fragment of ColI*drd-1* in pLG214 encodes two smaller polypeptides of *M*_r 87,000 and 42,000.

The 240,000 *M*_r protein was also detected in extracts of cells carrying pLG215 or ColI*drd-1* but not pLG214, pBR325 or ColI. Representative results are in Fig. 3, lanes e, g and h. The protein is the largest of the species that stain with Coomassie blue. Lane f contains an extract of BW84 (pLG216). BW84 is the *uvrB*⁺-*chlA*⁺ parent¹³ of BW86 and pLG216 is a *Sog*⁺ plasmid identified after *in vitro* mutagenesis of pLG211 with hydroxylamine¹⁷. The strain was fortuitously isolated during the screening of extracts of appropriate Ap^r BW84 transformants for primase activity. It has the advantage of producing about five times more plasmid primase activity than bacteria harbouring ColI*drd-1* or pLG215; it also overproduces the 240,000 *M*_r protein.

Purification of a DNA primase and antigenic relationship of polypeptides

A plasmid DNA primase was purified from BW86 (pLG214) using the procedure given in Fig. 4 legend. Chromatography on heparin-Sepharose was very useful because it allowed a 100-fold purification of enzyme in the first step starting with a crude cell extract. The preparation of DNA primase contained one major protein of *M*_r 87,000 which co-sedimented with enzyme activity through a glycerol gradient (Fig. 4). Its sedimentation coefficient was calculated to be 4.2S, indicating that it functions as a single subunit protein. This DNA primase is assumed to be the larger of the two polypeptides detected in pLG214-containing minicells (Fig. 3, lane c).

Using antiserum raised against this DNA primase preparation, immunological studies indicated an antigenic relationship between the two polypeptides encoded by the fragment of ColI*drd-1* present in pLG215 and the smaller pair of proteins specified by pLG214. The procedure, adapted from that of Towbin *et al.*¹⁸, involved separation of proteins in crude cell extracts by SDS-gel electrophoresis followed by their electrophoretic transfer to a nitrocellulose sheet. The resulting blot was incubated first with antiserum raised¹⁹ in a rabbit against the pLG214-specified primase and then with fluorescein-conjugated goat anti-rabbit immunoglobulins to detect bound antibody.

Figure 5, lanes c and d, show that antiserum reacted with two polypeptides of *M*_r 240,000 and 180,000 in extracts of bacteria harbouring pLG216 or pLG215. In the case of BW84

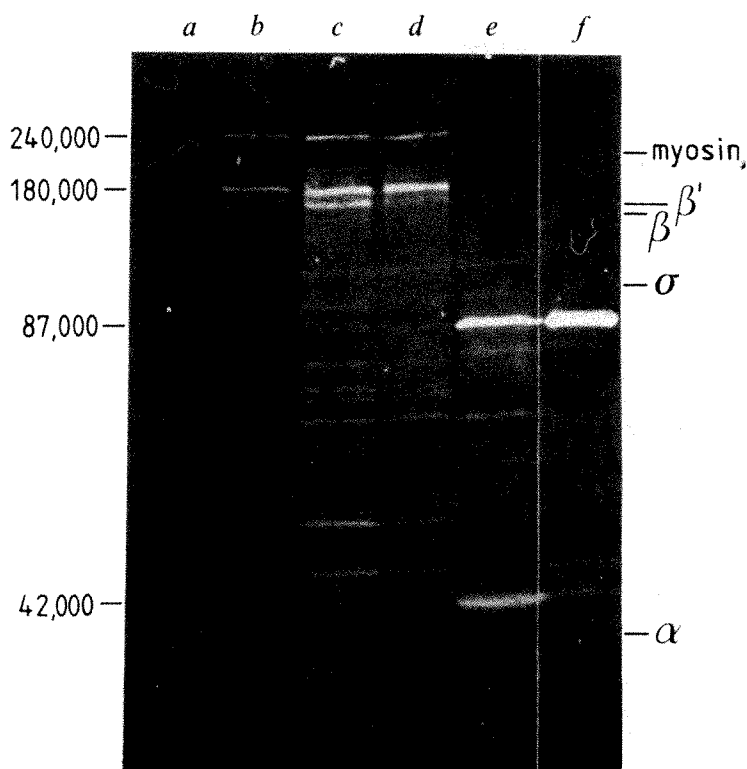


Fig. 5 Immunological analysis of proteins encoded by *Sog*⁺ plasmids. Crude cell extracts were prepared and separated on an SDS-polyacrylamide gel, as described in Fig. 3 legend. Proteins were transferred electrophoretically to a nitrocellulose membrane and reacted with antiserum as described elsewhere¹⁸. Antiserum against purified pLG214-specified DNA primase was applied in a 1:150 dilution. Fluorescein-conjugated goat anti-rabbit IgG were diluted 1:50. Lane a, BW86 (pBR325); b, BW86 (ColI*drd-1*); c, BW84 (pLG216); d, BW86 (pLG215); e, BW86 (pLG214); f, DNA primase purified from BW86 (pLG214). Migration of molecular weight standards, given under Fig. 3, is shown on the right margin.

(pLG216), the strain that overproduces plasmid primase, a third immunogen (*M*_r 168,000) was also apparent (lane c). The significance of this protein has not been established. Lane e shows that BW86 (pLG214) contains two smaller immunogens (*M*_r 87,000 and 42,000). Only the larger, assumed to be the DNA primase, was present in the antigen (lane f). Preimmune serum was unreactive. Two observations suggest that these polypeptides are encoded solely by the cloned fragment of ColI*drd-1* and are not products of fused cloned and vector DNA. First, antiserum did not react with protein in an extract of cells harbouring pBR325 (lane a). Second, the 240,000 and 180,000 *M*_r polypeptides are indistinguishable by size from a pair of polypeptides detected in bacteria containing the parental plasmid, ColI*drd-1* (lane b).

Antiserum, but not preimmune serum, inhibited the ability of crude extracts of BW86 containing either ColI*drd-1* or pLG215 to promote DNA synthesis on viral strands of phage fd (data not shown). Thus, the primase encoded by these plasmids is structurally related to the 87,000 *M*_r primase of pLG214, the antigen. Similarly, antiserum blocked activity of the primase specified by two other IncIα plasmids, R144*drd-3* and R64*drd-11*, and the IncIγ plasmid, R621a. In contrast, the antibodies failed to inhibit DNA synthesis on viral strands of phages ΦX174 and G4 in an ammonium sulphate fraction¹⁴ of *dna*⁺ cells (data not shown). Thus, the DNA primases encoded by IncIα plasmids are antigenically unrelated to the *E. coli* enzyme.

Products of the *sog* gene

The following hypothesis is proposed to explain the protein relationships. The *sog* gene encodes a pair of sequence-related

polypeptides of apparent molecular weight 240,000 and 180,000. The deletion mutation in pLG214 disrupts synthesis of the C-terminal segments giving a pair of truncated polypeptides (M_r 87,000 and 42,000). The smaller protein in each pair may be a processed product of the larger or may be generated from a second translational start within the *sog* gene without a change in reading frame. The latter possibility, for which gene A of Φ X174 is a precedent²⁰, is favoured by the properties of two *Sog*⁻, primase-defective, amber mutants of pLG211; these specify synthesis of the 180,000 M_r protein but not the 240,000 M_r protein (to be published). It is proposed that DNA primase activity requires part of the N-terminal sequence of the 240,000 and 87,000 M_r proteins. This is consistent with the properties of the amber mutants and the finding that the 87,000 M_r protein, but not the 42,000 M_r protein, purified as a DNA primase. Furthermore, primase activity *in vitro* has been detected in a partially purified preparation of the 240,000 M_r polypeptide (in preparation).

Suppression and DNA primase

The deletion in pLG214 enhances the efficiency of suppression of the *dnaG3* mutation and generates a truncated protein that purifies as a DNA primase. We therefore conclude that plasmid primase is the enzyme responsible for suppression. The enzyme is distinguishable from *dnaG* protein on the basis of molecular weight, antigenic determinants and product of reaction; *in vitro* it directs synthesis of a primer that contains cytidine or cytidine 5'-monophosphate at the 5' terminus (E.L., in preparation) and not an adenosine nucleotide as in the primer synthesized by *dnaG* protein. Thus, the phenomenon of suppression reflects the substitution of an unrelated plasmid enzyme for *E. coli* primase in the discontinuous synthesis of bacterial DNA and any primase-dependent event that may occur at the origin of DNA

replication. The plasmid primase can also substitute *in vivo* during replication of phage λ DNA (unpublished data) and of ColEI-type replicons like the recombinant plasmids used here. Thus DNA primases encoded by Inc α plasmids are remarkable for their ability to use a large variety of templates¹⁴. IncP plasmids, represented by RP4, also specify a DNA primase active in the assay reported here (E.L., unpublished data). This enzyme may also have the potential to act in bacterial DNA replication because derivatives of another IncP plasmid, R68-45, can suppress the *dnaG3* mutation²¹.

Inc α plasmid primase seems to cooperate with DNA polymerase III holoenzyme. Extension of primers *in vitro* requires DNA polymerase III and the *dnaZ* protein¹⁴, which is a component of the holoenzyme²². *In vivo*, bacterial DNA replication in the *dnaZ* mutant, AX727, remains temperature sensitive in the presence of ColIdrd-1 (unpublished data).

ColIdrd-1 and R64drd-11 also fail to suppress the temperature sensitivity of DNA replication in *dnaB70* and *dnaB43* mutants^{23,24}, implying that plasmid primase cannot bypass the requirement for *dnaB* protein. The *dnaB* product is thought to act in discontinuous DNA synthesis as a 'mobile promoter' for primer synthesis by bacterial primase⁵. Unlike the latter enzyme, R64drd-11 primase acts independently of *dnaB* protein in the priming reaction on single-stranded DNA of phage Φ X174¹⁴. Thus, the requirement for *dnaB* protein in plasmid primase-mediated DNA replication may reflect involvement of the protein in some process other than the priming reaction—possibly, unwinding of the DNA helix²⁵.

We thank Georg Stöffler for preparing antiserum, Brigitte Geschke and Marianne Schlicht for assistance, Pam Kramer for the electron microscopy, and Alec Jeffreys for his interest. Part of this work was supported by an EMBO short-term fellowship to B.M.W. and SRC grant GR/B24172. We thank H. Schuster for encouraging the collaboration.

Received 4 November 1980; accepted 20 January 1981.

- Ogawa, T. & Okazaki, T. A. *Rev. Biochem.* **49**, 421–457 (1980).
- Lark, K. G. *Nature new Biol.* **240**, 237–240 (1972).
- Rowen, L. & Kornberg, A. *J. biol. Chem.* **253**, 758–764 (1978).
- Wickner, S., Wright, M. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1613–1618 (1973).
- McMacken, R., Ueda, K. & Kornberg, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4190–4194 (1977).
- Scherzinger, E., Lanka, E., Morelli, G., Seiffert, D. & Yuki, A. *Eur. J. Biochem.* **72**, 543–558 (1977).
- Richardson, C. C. *et al. Cold Spring Harb. Symp. quant. Biol.* **43**, 427–440 (1978).
- Nossal, N. J. *J. biol. Chem.* **255**, 2176–2182 (1980).
- Liu, C. C., Burke, R. L., Hibner, U., Barry, J. & Alberts, B. M. *Cold Spring Harb. Symp. quant. Biol.* **43**, 469–487 (1978).
- Jacob, A. E. *et al. in DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 607–638 (Cold Spring Harbor Laboratory, New York, 1977).
- Wilkins, B. M. *J. Bact.* **122**, 899–904 (1975).
- Sasakawa, C. & Yoshikawa, M. *J. Bact.* **133**, 485–491 (1978).
- Boulnois, G. J. & Wilkins, B. M. *Molec. gen. Genet.* **175**, 275–279 (1979).
- Lanka, E., Scherzinger, E., Günther, E. & Schuster, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3632–3636 (1979).
- Bolivar, F. *Gene* **4**, 121–136 (1978).
- Brutlag, D., Schekman, R. & Kornberg, A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2826–2829 (1971).
- Eichenlaub, R. *J. Bact.* **138**, 559–566 (1979).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
- Stöffler, G. & Wittmann, H. G. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2283–2287 (1971).
- Linney, E. & Hayashi, M. *Nature new Biol.* **245**, 6–8 (1973).
- Ludwig, R. A. & Johansen, E. *Plasmid* **3**, 359–361 (1980).
- McHenry, C. & Kornberg, A. *J. biol. Chem.* **252**, 6478–6484 (1977).
- Boulnois, G. J., Beddoes, M. J. & Wilkins, B. M. *J. Bact.* **138**, 324–332 (1979).
- Fenwick, R. G. Jr & Curtiss, R. III *J. Bact.* **116**, 1212–1223 (1973).
- Staudenbauer, W. L., Scherzinger, E. & Lanka, E. *Molec. gen. Genet.* **177**, 113–120 (1979).
- Godson, G. N. & Sinsheimer, R. L. *Biochim. biophys. Acta* **149**, 476–488 (1967).
- Sutcliffe, J. G. *Cold Spring Harb. Symp. quant. Biol.* **43**, 77–90 (1978).
- Jeffreys, A. J. & Flavell, R. A. *Cell* **12**, 429–439 (1977).
- Jeffreys, A. J. *Cell* **18**, 1–10 (1979).
- Daniels, D. L., de Wet, J. R. & Blattner, F. R. *J. Virol.* **30**, 390–400 (1980).
- Sutcliffe, J. G. *Nucleic Acids Res.* **5**, 2721–2728 (1978).
- Timmis, K. N., Cabello, F. & Cohen, S. N. *Molec. gen. Genet.* **162**, 121–137 (1978).
- Davis, R. & Vapnek, D. *J. Bact.* **125**, 1148–1155 (1976).
- Humphries, P. *et al. Nucleic Acids Res.* **5**, 905–924 (1978).
- Achtman, M., Manning, P., Edelbluth, C. & Herrlich, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4837–4841 (1979).
- Dougan, G. & Sherratt, D. *Molec. gen. Genet.* **151**, 151–160 (1977).
- Lanka, E., Edelbluth, C., Schlicht, M. & Schuster, H. *J. biol. Chem.* **253**, 5847–5851 (1978).
- Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83–88 (1974).

The *E. coli* β -lactamase attenuator mediates growth rate-dependent regulation

Bengtåke Jaurin, Thomas Grundström, Thomas Edlund & Staffan Normark

Department of Microbiology, University of Umeå, S-901 87 Umeå, Sweden

We have identified a new control or attenuator region in the chromosomal β -lactamase operon of Escherichia coli. A single base alteration within this attenuator led to a loss in the cell's ability to coordinate its content of β -lactamase with growth rate. We suggest a mechanism through which this mode of regulation operates.

MOST of the more abundant *E. coli* proteins increase in specific amounts with increasing growth rates¹. This group of proteins includes ribosomal proteins, elongation factors, aminoacyl

tRNA synthetases and subunits of RNA polymerase. Also, the amount of rRNA and tRNA increases coordinately with growth rate. Both active and passive regulatory models have been

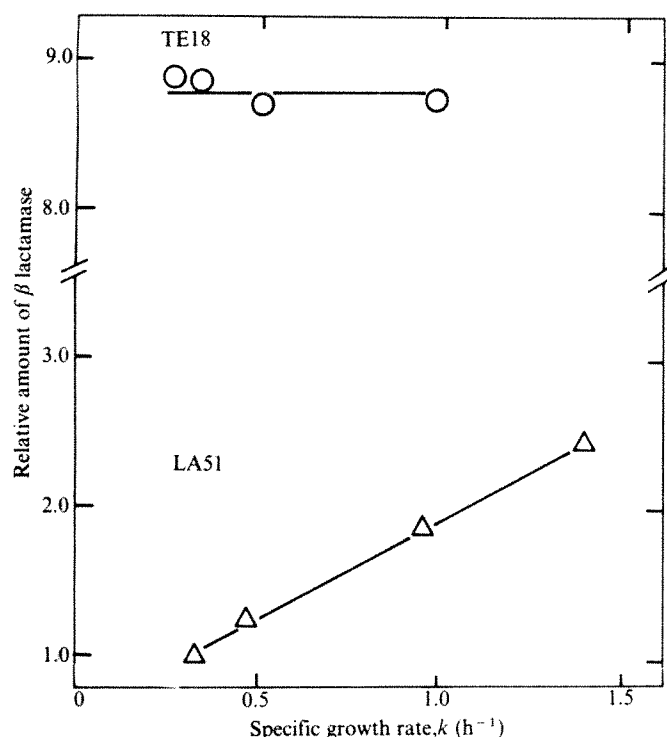


Fig. 1 Relative amount of β -lactamase at different growth rates at 37 °C in *E. coli* K12 strains LA51 (*ampP15G16*) and TE18 (*ampP15G16*, *ampL35A*). The relative amount of β -lactamase per total protein was determined with rocket immunoelectrophoresis as previously described³ and plotted as a function of the first-order constant for growth (k), as calculated from the expression $k = \ln 2/\text{mass doubling time in hours}$. The media used were as follows: medium E + glycerol, medium E + glucose, medium E + glucose + casamino acids, and LB medium³.

proposed to explain the positive correlation between the production of these components of the translation machinery and growth rate².

The relative amount of chromosomally encoded β -lactamase of *E. coli* K12 also increases with growth rate³. This periplasmic enzyme hydrolyses the β -lactam ring of both cephalosporins and penicillins including ampicillin. Its structural gene, *ampC*, has been located at 93.8 min on the *E. coli* linkage map^{4,5}. The strict linear relationship between β -lactamase production and ampicillin resistance enables one to isolate mutants with altered production of β -lactamase by selecting on plates containing different concentrations of ampicillin^{4,6}.

Ampicillin-resistant mutants have been found to carry either *cis*-dominant mutations in a control sequence region denoted *ampA* that is adjacent to *ampC*⁶, or multiple tandem repeats of DNA segments encoding the structural gene⁷. The regulatory region *ampA*, as well as the structural gene *ampC*, have been cloned on a ColE1 hybrid plasmid and by further subcloning localized to a 1,370-base pair DNA segment⁶.

We describe here the sequence and structure of the control region for the β -lactamase gene, *ampC*, in wild-type *E. coli* and two regulatory mutants, and suggest a mechanism whereby the control region regulates transcription in response to the quantity of ribosomes per cell.

A mutation in the control region abolishes the positive correlation between β -lactamase production and growth rate

Strain LA5 is wild type for β -lactamase production. The concentration of ampicillin on broth plates required to kill 50% of the applied wild-type cells (LD_{50}) is $1 \mu\text{g ml}^{-1}$. Strain LA51 is an ampicillin-resistant transductant of strain LA5 (ref. 4). It shows an LD_{50} value of $15 \mu\text{g ml}^{-1}$ of ampicillin, and a corresponding 15-fold increase in β -lactamase production. This strain carries a *cis*-dominant regulatory mutation denoted *ampA1* (ref. 4), which we have recently localized within a

370-base pair DNA segment adjacent to *ampC*⁶. From strain LA51, spontaneous mutants with a further increase in ampicillin resistance have been isolated⁶, one of which, TE18, exhibiting an ampicillin resistance of $62 \mu\text{g ml}^{-1}$ (LD_{50}), carries a second mutation in the same 370-base pair regulatory region for *ampC*.

We have previously demonstrated that the specific amounts of β -lactamase produced both in the wild-type strain LA5 and in the mutant strain LA51 increase proportionally with growth rate³. In contrast, when the growth rate responses of the mutant strains LA51 and TE18 were compared (Fig. 1) it was evident that the growth rate-dependent regulation of β -lactamase had been abolished in the double mutant TE18. Thus, the level of β -lactamase in strain TE18 was much higher than that of LA51,

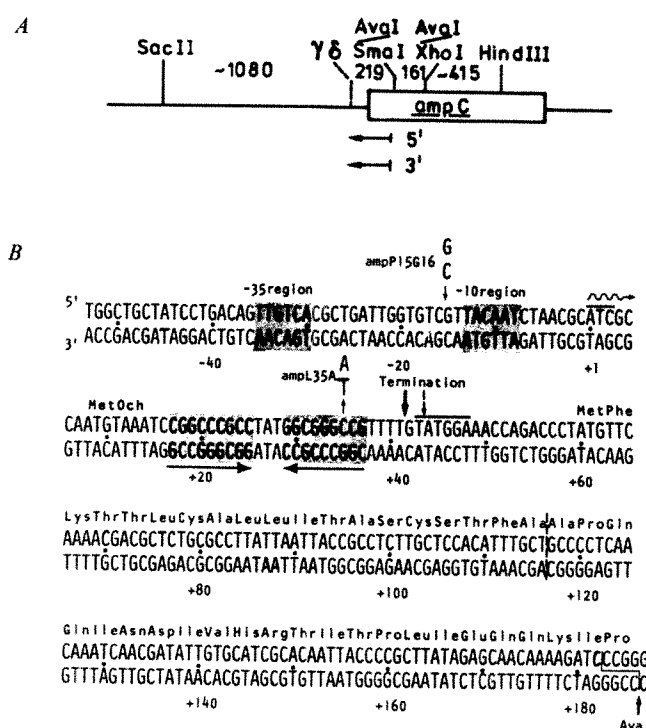


Fig. 2 A, restriction map of the *amp* region and sequencing strategy. The dashed line designates the insertion point of the transposable element $\gamma\delta$ (ref. 6). The box indicates the position of the structural gene, *ampC*. The direction of sequencing from the 5' and 3' ends of DNA cut with *AvaI* (*SmaI*) is indicated by arrows. The 1.5-kilobase *XhoI*-*SacII* fragments from plasmids pNU6, pNU28 and pNU35 were prepared from polyacrylamide gels²¹ and cleaved with *AvaI*. For the 5'-end labelling, 5–30 pmol of 5' ends were incubated in 10 μl of 50 mM Tris-HCl, pH 8.0, with 0.03 units of bacterial alkaline phosphatase at 45 °C for 30 min. 1 μl of 50 mM EGTA was added and the solution was heated at 65 °C for 15 min. 150 μCi vacuum-dried [γ -³²P]ATP ($>5,000 \text{ Ci mmol}^{-1}$, Radiochemicals) was dissolved in 9 μl of H_2O and transferred to the DNA mixture. 2.5 μl of 0.5 M Tris-HCl, pH 7.6, 0.1 M MgCl_2 , 50 mM dithiothreitol (DTT), 1 mM spermidine, 1 mM EDTA were added and the reaction mixture incubated with 5 units of T4 polynucleotide kinase (Biolabs) at 37 °C for 25 min. After one phenol extraction, labelled DNA was isolated by gel filtration on a Sephadex G-100 column. The 3'-end labelling was in principle as described by Schwartz *et al.*²². However, [α -³²P]dCTP (Radiochemicals) was used, followed by a chase for 10 min at 25 °C with 0.05 mM of cold dCTP. The labelled fragments were cleaved with *Bst*NI (cleaves 625 base pairs to the right of the *SacII* cut). The *AvaI*-*Bst*NI fragments were purified from polyacrylamide gels²³ and subjected to the nucleotide sequencing procedure described by Maxam and Gilbert²³. The degradation products were separated on 0.5-mm thick 8% and 20% sequencing gels with subsequent autoradiography at -20°C . B, nucleotide sequence of the control region and the N-terminal part of *ampC*. Start of transcription is marked by +1 and the wavy arrow. The major and minor termination points of the short transcript are indicated by vertical solid and dashed arrows, respectively. The stem and loop structure is marked by horizontal arrows and stippled boxes. Solid lines designate possible base pairing with the 3' end of 16S RNA¹⁸. Met and Och designate initiation codon (methionine) and stop codon (ochre), respectively. The limit between the presumed signal peptide and the mature β -lactamase is marked by the vertical dashed line. The positions of the G-C insertion in *ampP15G16* and the A-T transversion in *ampL35A* are represented by vertical arrows. The -35 and -10 promoter regions are marked by stippled boxes.

and it did not change when the steady-state growth rate was varied by a factor of three.

Establishment of the DNA sequences in the wild-type and regulatory mutants

Plasmid pNU6 is a pBR322 derivative which carries the *ampC* gene and the wild-type regulatory region. Plasmids pNU28 and pNU35 carry *ampC* and the mutated control regions of strains LA51 and TE18, respectively⁶. Figure 2A shows a restriction map of a 2.5-kilobase DNA segment containing the β -lactamase gene. We have previously located the regulatory mutations of LA51 and TE18 between the *XhoI* site and the insertion point for the transposable element $\gamma\delta$ (ref. 6) as shown in Fig. 2A. The nucleotide sequences from the *AvaI* (*SmaI*) site towards the *SacII* site were obtained by both 5'- and 3'-end-labelling and are presented in Fig. 2B. This stretch of DNA was sequenced from plasmids pNU6, pNU28 and pNU35. To localize the beginning of the coding region for *ampC*, the order of the 12 N-terminal amino acids of purified chromosomal β -lactamase was determined by Edman degradation (manuscript in preparation). There was perfect correspondence with the codons from base pairs +117 to +152. The nearest predicted translation start sequence appeared 19 codons before that for the N-terminal amino acid (alanine) of the purified enzyme. These data are consistent with the interpretation that the translational product of *ampC* carries a 19-amino acid N-terminal extension. This stretch of amino acids has all the common structural features of a signal peptide⁸.

Upstream of the transcriptional start base (+1, see below), sequences at nucleotide positions -13 to -8 (-TACAAT-) and -35 to -30 (-TTGTCA-) both show extensive homologies with the conserved -10 and -35 regions of prokaryotic promoters⁹. Also, the sequences between and surrounding these regions show homology with known promoter sequences. When the DNA sequence of the same segment as in pNU6 was determined with the mutant plasmid pNU28, the only difference from the wild type was an insertion of a G·C base pair between positions -16 and -15 (Fig. 2B). Thus, this mutation, denoted *ampP15G16* (previously *ampA1*), which causes a 15-fold increase in β -lactamase synthesis, was located in the presumed *amp* promoter between the conserved -35 and -10 regions.

In the wild-type plasmid pNU6 and in plasmid pNU28, we find a 9-base pair long, exclusively G↔C-containing, dyad symmetry sequence at nucleotide positions +17 to +25 and +29 to +37. This symmetrical DNA sequence is followed by a stretch of four T residues on the noncoding strand. Many terminators for the *E. coli* RNA polymerase have similar sequence features¹⁰. Plasmid pNU35, carrying the *ampC* region of strain TE18, contained in addition to the mutation *ampP15G16*, a C·G to A·T transversion at position +35 (Fig. 2B). This *amp* leader mutation, denoted *ampL35A*, has occurred within the postulated terminator sequence.

In vitro transcription from DNA segments carrying a wild-type or a mutated regulatory region

The 1.5-kilobase *SacII*-*XhoI* DNA fragment which carries the entire sequenced region shown in Fig. 2B, was isolated from plasmids pNU6, pNU28 and pNU35 and added individually to an *in vitro* transcription system. The results are shown in Fig. 3A. From the pNU6 fragment that carries the wild-type regulatory sequence, both a long and a short [α -³²P]CTP-labelled transcript were obtained (L and S in Fig. 3A). The distance from the initiation site for transcription (see below) to the *XhoI* end of the *SacII*-*XhoI* fragment is ~340 base pairs. The long transcript is approximately this length and therefore probably represents a run-off transcript. Furthermore, when the DNA fragment was cleaved at the *AvaI* (*SmaI*) site located 161 base pairs from the *XhoI* site, the 340-base pair long transcript was not found.

Instead, as expected, a new transcript (~180 bases long) appeared (data not shown). The short transcript (S in Fig. 3A) was shown by RNA sequencing analysis to be 41 bases long (see below). The DNA sequence (Fig. 2B and data not shown)

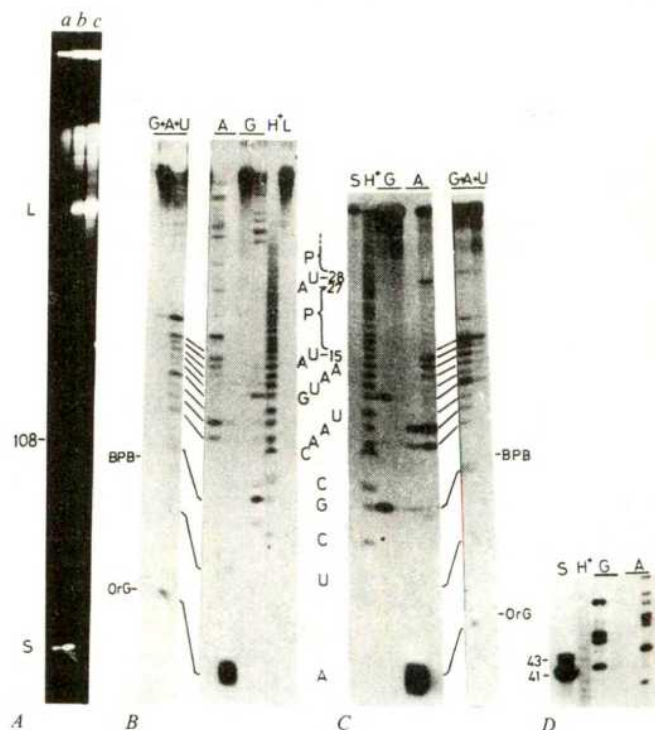


Fig. 3 In vitro transcription and RNA sequencing of the *amp* leader region.

A shows transcripts obtained from the 1.5-kilobase *XhoI*-*SacII* fragments (see Fig. 2A) carrying different *amp* alleles. The *in vitro* transcription reactions contained 0.25 μ g of the DNA fragment, 2 units of RNA polymerase (Boehringer Mannheim), 100 μ M each of GTP, UTP and ATP, 10 μ M [α -³²P]CTP (1.5 Ci mmol⁻¹) and RNA polymerase binding buffer (5% glycerol, 20 mM Tris-HCl pH 8.0, 10 mM MgCl₂, 0.1 mM EDTA, 0.1 mM DTT and 120 mM KCl) in a total volume of 20 μ l. After 10 min incubation at 37°C, 8 μ l of formamide-buffer electrophoresis dyes²³ supplemented with 0.4 μ g tRNA were added, and the samples were lyophilized to 8 μ l volume. A 6% acrylamide: bis (30:1) gel with 8 M urea was run²³ and autoradiographed overnight. The position of a 108-base transcript from ColE1²⁴ is indicated. L and S indicate long (340 bases) and short (41 bases) transcripts, respectively. The sources of the DNA fragments used were pNU6 (*ampP*⁺, *ampL*⁺) (lane a), pNU28 (*ampP15G16*, *ampL*⁺) (lane b) and pNU35 (*ampP15G16*, *ampL35A*) (lane c). B and C show autoradiographs of RNA sequencing gels of the long and short transcripts, respectively. The 5'-end-labelled transcripts were synthesized in RNA polymerase binding buffer supplemented with 150 μ M CTP and UTP, 40 μ M GTP, 40 μ M [γ -³²P]ATP (100 Ci mmol⁻¹) and 10 units RNA polymerase in a total volume of 100 μ l. The long transcript was synthesized with 2.2 μ g of the 1.3-kilobase *SmaI*-*SacII* fragment from pNU35 as template, and the short transcript with 1.25 μ g of the 1.5-kilobase *XhoI*-*SacII* fragment from pNU28. After incubation at 37°C for 15 min, the samples were treated as above except that the gel contained 8% acrylamide. The transcripts were isolated by the method of Maxam and Gilbert²⁵ and the RNA sequencing reactions were as described by Krupp and Gross²⁵. All reactions were done with 4 μ g of carrier tRNA. The G+A+U reactions were carried out with 5 and 50 μ g of *Neurospora* endonuclease, respectively, the A reactions with 0.05 and 0.5 units of RNase U₂, respectively, and the G reactions with 1 and 10 units of RNase T₁, respectively. H⁺ indicates controlled partial acid hydrolysis giving labelled mRNA of all sizes. L and S indicate untreated long and short transcripts, respectively. The position of the orange G and bromophenol blue markers which migrate with 2 and 6 bases length, respectively²⁶, on the 25% polyacrylamide-urea gels used is marked OrG and BPB. Note that due to the absence of a phosphate at their 3' ends, oligonucleotides produced by *Neurospora* nuclease move more slowly in the gel²⁵. These bands have been connected with the corresponding bands from the other reactions. The sequence of the transcripts is indicated between B and C. P indicates positions where the transcripts due to their secondary structure are protected from the action of the enzymes. D is an autoradiograph of a part of an 8% polyacrylamide gel, showing the sequence of the long transcript around the termination point of the short transcript. Note that in contrast to the transcripts, the labelled oligonucleotide degradation products have a phosphate at their 3' end. Consequently, when interpreting D it must be remembered that the untreated ladder is displaced upwards by about half a nucleotide unit. Most of the short transcripts terminate at base 41 of the long transcript with minor termination at base 43.

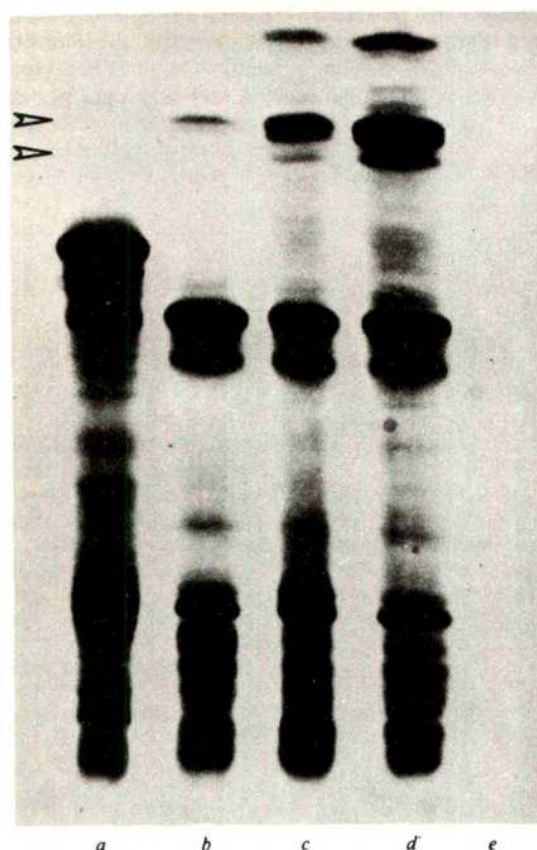


Fig. 4 Expression of β -lactamase in a coupled *in vitro* transcription-translation system. The system is originally derived from Zubay²⁷, and modified by Johnsen²⁸ and Collins²⁹. The polypeptides were separated on a 10–17.5% gradient polyacrylamide gel as previously described⁶. The gel was fluorographed using En³hance (NEN). The plasmids used were: a, pBR322; b, pNU6; c, pNU28; d, pNU35; e, no plasmid DNA added. The 36K and 38K proteins which represent mature β -lactamase and presumed precursor β -lactamase, respectively, are marked by arrows.

reveals that the long transcript contains seven times as many C residues as the short one. Taking this into account, and with the relative band intensities obtained by scanning of the autoradiographs, we conclude that the short transcript is reproducibly about 15 times more abundant than the longer one.

In vitro transcription with the same amount of the analogous DNA fragment from pNU28 (carrying the *ampP15G16* mutation) resulted in an increase in the synthesis of both the long and the short transcripts (Fig. 3A, lane b). This contrasts with the result obtained when the corresponding *SacII-XhoI* fragment from pNU35 (carrying both the *ampP15G16* and *ampL35A* mutations) (lane c) was transcribed; no synthesis of the short 41-base transcript was observed. Moreover, the amount of the long run-off transcript was increased. The loss of the short transcript associated with an increase in the amount of the long run-off transcript suggests that the second mutation in TE18, *ampL35A*, abolishes termination in the *amp* leader region.

Localization of initiation and termination points for transcription

The initiation nucleotide *in vitro* of β -lactamase mRNA was determined by transcription using 10 μ M of one nucleotide triphosphate (NTP) and 150 μ M of the other three NTPs. Synthesis of β -lactamase transcripts was severely inhibited by ATP limitation but not by limitation of any of the other three NTPs (data not shown). This suggests that A is the 5' nucleotide. The long run-off transcript and the short transcript were labelled

using [γ -³²P]ATP, and subjected to RNA sequencing by partial enzymatic degradation (Fig. 3B, C). The first 16 nucleotides in the two transcripts were identical and homologous to the DNA sequence from bases +1 to +16. This shows that the A at +1 is the start base for both these mRNAs and confirms the position for the β -lactamase promoter. The nucleotides in the RNA sequence between positions 17 and 37 except for the A and U residues at positions 27 and 28, respectively, were protected against enzymatic cleavage. As the enzymes used are specific for single-stranded RNA¹¹, this finding implicates the existence of a 9-base pair RNA stem with a 3-base loop at positions 26–28. The point of termination of the short transcript was determined by comparing the length of non-degraded short transcript with that of enzymatically degraded long transcript (Fig. 3D). At least 99% of the short transcript ended with the U at position 41. A minor fraction terminated two nucleotides later.

Expression of the *amp* operon in a coupled *in vitro* transcription-translation system

When DNA from plasmids pNU6, pNU28 and pNU35 were added to a coupled transcription-translation system a protein with an apparent molecular weight (MW) of 38,000 (38K) was produced. The relative production of this protein from the respective plasmids was ~1:15:60 (Fig. 4). Purified β -lactamase migrates with an apparent MW of 36K. A faint band appeared at this position in the gel when plasmid pNU6 was used. In the mutants the corresponding protein band was more intense. We suggest that the 38K protein observed is the precursor form of the *E. coli* chromosomal β -lactamase postulated on the basis of the DNA sequence. The uppermost band with an apparent MW of 56K also increased in intensity in the mutants—the nature of this band is not known.

The fourfold difference between plasmid pNU28 and pNU35 in production of the pre- β -lactamase in the coupled transcription-translation system, corresponds to the fourfold differences in steady-state levels of β -lactamase found *in vivo*. On the other hand, in the *in vitro* transcription system the difference in transcriptional readthrough was about 16-fold. This suggests that a factor(s) present *in vivo* and in the coupled

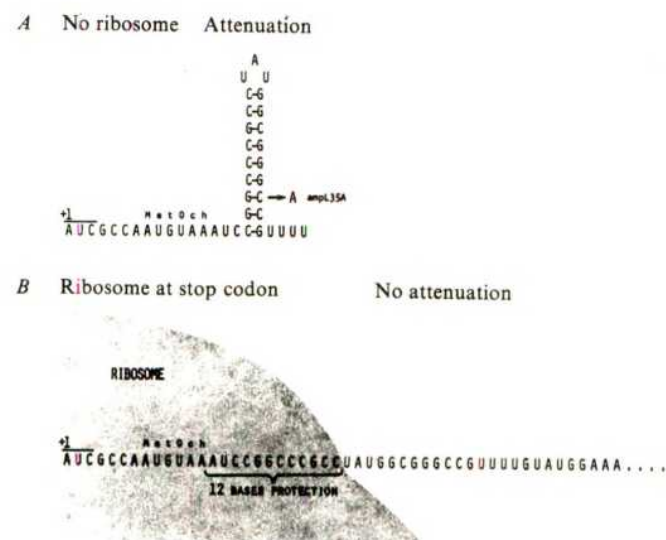


Fig. 5 A model for growth rate-dependent attenuation of transcription. The stippled area represents a ribosome. The symbols are the same as in Fig. 2B. A, terminated transcript; B, hypothetical protection by the ribosome preventing formation of the termination stem, leading to readthrough (see text for discussion).

system increases the transcriptional readthrough about fourfold compared with a pure *in vitro* transcription system.

Features of the β -lactamase attenuator

One of our major conclusions is that the *amp* operon contains within its leader region a site for termination of transcription that is recognized by the *E. coli* RNA polymerase *in vitro*. Factors such as the rho protein¹² are not required for efficient termination *in vitro* at this site. However, we do not know whether the rho protein has any effect in the *in vivo* situation. On the other hand, the terminator in the *amp* leader has a structure similar to those of sequenced rho-independent terminators¹⁰.

The base pair substitution *ampL35A* occurred 7 base pairs upstream from the major termination site and within the 9-base pair symmetrical DNA sequence (Fig. 2B). From both thermodynamic considerations and the relative susceptibility to enzymatic cleavage, we suggest that the leader RNA transcript forms a secondary structure with a 9-base pair G $\leftrightarrow$ C-containing stem and a 3-base loop (Fig. 5A). The *ampL35A* mutation decreases the thermodynamic stability of such an RNA stem from a ΔG value of -29.4 kcal to -17.4 kcal (ref. 13). Termination of transcription *in vitro* decreases to less than 1% in template with the *ampL35A* mutation. It seems likely that the terminator in the *amp* leader also operates *in vivo*, because β -lactamase synthesis at the maximal growth rate is roughly four times higher in the *ampL35A*-carrying mutant strain, TE18, than in its parent strain, LA51.

Growth rate-dependent regulation of *ampC* is exerted at the level of anti-termination

In the *trp*, *phe*, *his*, *leu*, *ilv* and *thr* biosynthetic operons, a potential coding region for a short leader peptide precedes a terminator¹⁴⁻¹⁶. In each case, the leader peptide contains 'regulatory' codons for the amino acids that regulate the operon. In the model originally put forward by Lee and Yanofsky¹⁷ it is proposed that synthesis of the leader peptide is involved in the regulation of transcription termination at the attenuator. This regulation depends on the formation of mutually exclusive RNA secondary structures. Ribosomes translating the entire coding region would permit formation of the terminator secondary stem and loop structure, whereas ribosomes slowed down at the 'regulatory' codons, due to a deficiency in the concentration of the charged cognate tRNA, would favour formation of an alternative structure and thereby preclude formation of the terminator structure.

The 41-base long *amp* leader transcript is considerably shorter than other known leader transcripts and has no potential coding region for a leader peptide. Moreover, the *amp* leader RNA cannot form any alternative secondary structure that precludes formation of the terminator stem and loop. However, examination of the RNA sequence reveals that the first three bases are complementary to a sequence near the 3' end of 16S rRNA (ref. 18). Furthermore, bases 8-13 consist of an initiation codon (AUG) that is followed directly by an ochre stop codon (UAA). Thus, the *amp* leader RNA has a potential ribosome binding site and may form a translation initiation complex. Our major concerns are whether or not this potential ribosome binding site is used and whether it could have a regulatory effect on the termination of transcription. A ribosome stalled at the stop codon would protect ~ 12 bases downstream¹⁹. As shown in Fig. 5B this could prevent formation of the termination stem and loop structure.

Expression from the *amp* operon is positively correlated with the steady-state growth rate. In theory, this growth rate-dependent regulation could occur at the level of initiation or

termination of transcription or both. The *ampP15G16* mutation increases transcription initiation *in vitro* by a factor of 15 without significantly affecting the degree of transcription termination. As this up-promotor mutation did not affect the growth rate-dependent expression of *ampC* we find it unlikely that the *amp* promoter participates in this regulation.

In the double mutant, TE18, expression of *ampC* has become independent of the growth rate. This fact, together with the observation that the mutation *ampL35A* abolishes termination of transcription *in vitro* at the *amp* attenuator, suggests that the growth rate-dependent regulation of the *amp* operon is exerted at the level of anti-termination. The anti-termination factor in our system is not known. We hypothesize that ribosome binding at the *amp* leader RNA has a regulatory role. The number of ribosomes per cell mass increases, over a broad range, proportionally with growth rate². If binding of ribosomes to the *amp* leader region is proportional to the cellular concentration of ribosomes, the degree of anti-termination would increase proportionally with growth rate.

Many proteins exhibit the same growth rate-dependent regulation as the chromosomal β -lactamase. For many operons this mode of regulation might also be exerted at the level of anti-termination. Terminator-like sequences have been found in the leader DNA of the *rpsE* and *rpsL* operons. From the *rpsL* operon, a short transcript has been obtained²⁰ which suggests that termination can at least occur *in vitro*. It is, however, not known if these possible terminator sequences are involved in the regulation of the operons.

The regulatory region of *ampC* consists of two control regions, a promoter and an attenuator. The up-promotor mutation described here changes the distance between the two conserved regions in the promoter. Possibly this is of importance for efficient RNA polymerase recognition. Because of the ease with which one can isolate and analyse mutations in the regulatory regions of the *amp* operon, this system can be used to increase our understanding of the functions of promoters and attenuators.

We thank Torbjörn Nilsson and Per Pettersson for help with the amino acid sequence determination, Monica Persson for technical assistance, and Stewart Cole, John K. Davies and Christopher Korch for careful reading of the manuscript. This work was supported by grants from the Swedish Natural Science Research Council (Dnr 3373) and from the Swedish MRC (Dnr 5428).

Received 28 July; accepted 24 December 1980.

- Pedersen, S., Block, P. L., Reeh, S. & Neidhardt, F. C. *Cell* **14**, 179-190 (1978).
- Maaløe, O. in *Biological Regulation and Development* Vol. 1 (ed. Goldberger, R. F.) 487-542 (Plenum, New York, 1979).
- Jaurin, B. & Normark, S. *J. Bact.* **138**, 896-902 (1979).
- Normark, S. & Burman, L. G. *J. Bact.* **132**, 1-7 (1977).
- Bachmann, B. J. & Low, B. K. *Microbiol. Rev.* **44**, 1-56 (1980).
- Grundström, T., Jaurin, B., Edlund, T. & Normark, S. *J. Bact.* **143**, 1127-1134 (1980).
- Edlund, T., Grundström, T. & Normark, S. *Molec. gen. Genet.* **173**, 115-125 (1979).
- Inouye, M. & Halegoua, S. *CRC Crit. Rev. Biochem.* **7**, 339-371 (1980).
- Siebenlist, U., Simpson, R. B. & Gilbert, W. *Cell* **20**, 269-281 (1980).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319-353 (1979).
- Peattie, D. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1760-1764 (1979).
- Pribnow, D. in *Biological Regulation and Development* Vol. 1 (ed. Goldberger, R. F.) 219-277 (Plenum, New York, 1979).
- Tinoco, I. Jr et al. *Nature new Biol.* **246**, 40-41 (1973).
- Keller, E. & Calvo, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6186-6190 (1979).
- Nargang, F. E., Subrahmanyam, C. S. & Umberger, H. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1823-1827 (1980).
- Lawther, R. P. & Hatfield, G. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1862-1866 (1980).
- Lee, F. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4365-4369 (1977).
- Shine, J. & Dalgarno, L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1342-1346 (1974).
- Steitz, J. in *Biological Regulation and Development* Vol. 1 (ed. Goldberger, R. F.) 349-399 (Plenum, New York, 1979).
- Post, L. E., Arfsten, A. E., Reusser, F. & Nomura, M. *Cell* **15**, 215-229 (1978).
- Akujärvi, G. & Pettersson, U. *Virology* **91**, 477-480 (1978).
- Schwartz, E., Scherer, G., Hobom, G. & Kössel, H. *Nature* **272**, 410-414 (1978).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Morita, M. & Oka, A. *Eur. J. Biochem.* **97**, 435-443 (1979).
- Krupp, G. & Gross, H. J. *Nucleic Acids Res.* **6**, 3481-3490 (1979).
- Horii, T., Ogawa, T. & Ogawa, H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 313-317 (1980).
- Zubay, G. *Rev. Genet.* **7**, 267-287 (1973).
- Johnsen, M. thesis, Univ. Copenhagen (1976).
- Collins, J. *Gene* **6**, 29-42 (1979).

LETTERS

IUE detection of bursts of H Ly α emission from Saturn

J. T. Clarke*, H. W. Moos*§, S. K. Atreya† & A. L. Lane‡

* The Johns Hopkins University, Baltimore, Maryland 21218, USA

† University of Michigan, Ann Arbor, Michigan 48109, USA

‡ Jet Propulsion Laboratory, Pasadena, California 91103, USA

An extended source of H Ly α emission (1,216 Å) from the vicinity of Saturn has been detected previously by a rocket-borne spectrometer¹, the OAO-Copernicus satellite², and the Pioneer 11 spacecraft³. During the rocket experiment, in which Saturn was observed on 15 March 1975, 700 R $\pm$ 50% of Ly α emission was detected from the planet and an additional 200 R $\pm$ 50% from a 53 arc s diameter circular area centred on the planet. This additional emission was believed to emanate from the ring system. The Copernicus observations, carried out in April 1976 and April 1977, yielded Ly α intensities of 1.40 $\pm$ 0.45 kR for the saturnian disk, <100 R for the rings, and 200 $\pm$ 100 R for Titan's torus. During the Pioneer 11 Saturn flyby in August–September 1979, the long-wavelength channel UV photometer detected emissions from the planet, the vicinity of the rings, and the orbital path of Titan; these emissions were interpreted as resonant scattering of solar Ly α radiation by H atoms and possibly auroral Ly α emission from the planet. Several mechanisms for producing H atoms outside of Saturn's atmosphere have been proposed, including sputtering of water ice from the rings by magnetospheric charged particles⁴, 'photo-sputtering' of the ring water ice by solar UV radiation⁵, and escape of H atoms from Titan's atmosphere⁶. A more recent suggestion is the possible impact on the rings and subsequent neutralization of the protons from an extended saturnian ionosphere¹⁵. We report here a new investigation of these potential sources of Ly α emission in a series of observations of the saturnian system carried out between January and July 1980 using the short wavelength (SWP) spectrograph of the IUE Observatory^{7,8}. North–south maps of the Ly α emission across the planet disk have shown pronounced spatial asymmetries in emission brightness. These asymmetries vary markedly on a time scale of days and are interpreted as bursts of Ly α emission of as much as 1 kR brightness averaged over a 6 $\times$ 10 arc s area, above a constant planetary emission level of 700–800 R. In fact, the Ly α emission peaks appear as essentially point source features in these data: if the emitting region is smaller than the 6 $\times$ 10 arc s instrumental resolution, the surface brightness must be proportionally higher.

Spatial imaging with the IUE has been accomplished both in a series of exposures of different regions of the saturnian system and by sampling within the field-of-view of the large spectrograph entrance aperture in single exposures. The image of the field-of-view subtended by the aperture is focused directly onto the detector at each wavelength; because the angular resolution of the instrument is 5–6 arc s on planetary objects and the large aperture extends 10 $\times$ 23 arc s, imaging perpendicular to dispersion is possible in a single exposure. By moving the $\sim$ 16 arc s diameter image of Saturn 5–10 arc s between exposures along the major axis of the aperture (which is constrained to point roughly north–south), maps of the disk emission plus the background geocoronal and interplanetary Ly α emission immediately north and south of Saturn have been obtained with 6 arc s spatial resolution. In each of the exposures centred on the polar regions, therefore, the level of background Ly α is directly measured and subtracted from the planetary emission; in exposures centred on Saturn the background level has been

determined by matching the level of planetary emission to that observed in the polar exposures immediately preceding or following. A relative sensitivity correction in the direction perpendicular to dispersion has been derived from exposures of the background Ly α emission and applied to the data reported here. The error in applying this correction has been determined by testing it on other exposures of background Ly α emission and measuring the average deviation from the mean level of emission; this error is $\pm$ 200 R. The pointing of the spacecraft during the north–south maps has been determined by noting the position in the aperture of the sharp drop in grating-scattered long wavelength radiation at the northern and southern edges of the saturnian disk. The north–south positioning of the aperture is believed to be accurate to within 1–2 arc s. This same technique of north–south mapping has been used in observations of the polar auroral Ly α emission from Jupiter⁹.

Representative north–south maps of the Ly α emission are plotted in Fig. 1 from observations performed on 19 January, 5 May and 22 July (all 1980). Non-equatorial peaks in the emission are apparent in the data from the first two days, while the July data are symmetric north–south with an equatorial brightness of 800 R. This equatorial brightness in the individual exposures on 22 July was measured over one-half of a rotation of the planet; therefore, 700–800 R is interpreted as a constant level of planetary emission from Saturn's upper atmosphere. The data of 22 July provide a good fit to a model emission profile

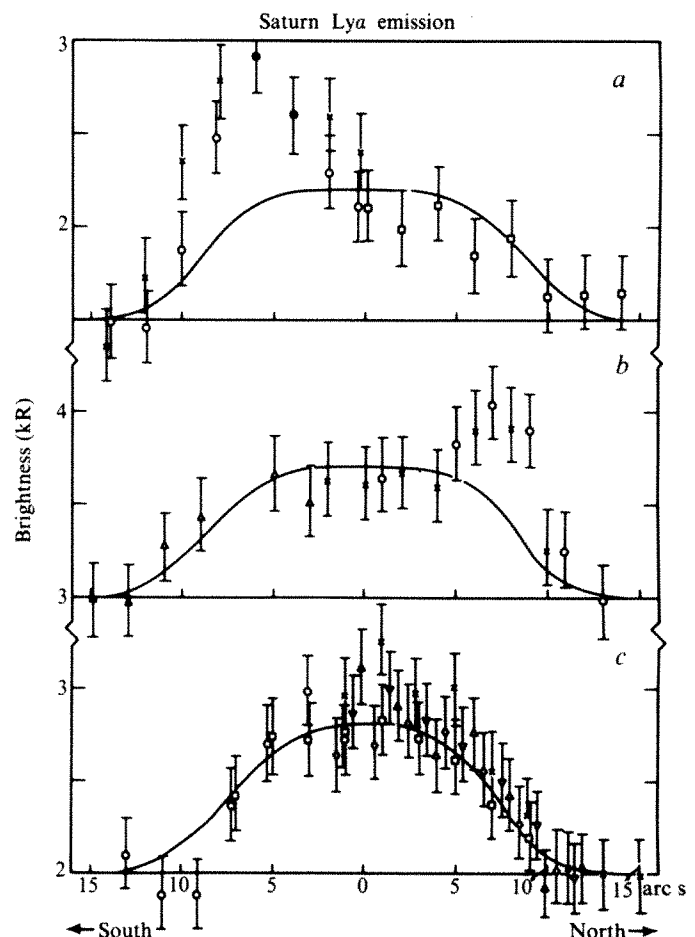


Fig. 1 North–south maps of the H Ly α emission from Saturn plus sky background. Different point characters represent different camera exposures, and the $\pm$ 200 R error bars are discussed in the text. The measured background varied by as much as $\pm$ 500 R in a series of exposures, and the exposures plotted have been normalized to the same level of background (set at the average background level observed on a given day). *a*, 19 January 1980; *b*, 5 May 1980; *c*, 22 July 1980.

§ Permanent address: Joint Institute for Laboratory Astrophysics and Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, Colorado, USA.

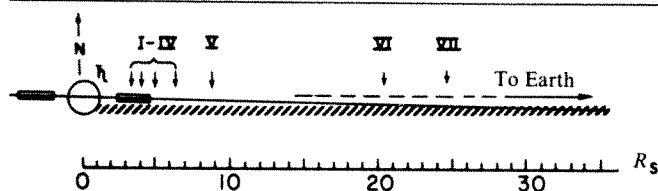


Fig. 2 Observing geometry during the exposures of 19 January 1980. The dashed line and arrow indicate the line-of-sight from the Earth, and the solid line represents the ring plane (which is tilted 1.6° south of the line-of-sight). The enhanced $\text{Ly}\alpha$ emission observed apparently from the south pole of Saturn could actually have come from anywhere in the shaded region. Roman numerals indicate the orbital radii of Saturn's inner moons (VI is Titan).

(the solid line in Fig. 1) representing a uniformly emitting disk of 800 R brightness smoothed by the 6 arc s instrumental resolution. Note that the decrease in emission from the equator to the poles in the model profile is due to the spatial resolution of the instrument, not limb darkening. If this emission is completely produced by resonant scattering of solar $\text{Ly}\alpha$ from upper atmospheric H it corresponds to an H column abundance on the order of 10^{14} cm^{-2} , following the simple resonant scattering model for Jupiter outlined by Clarke *et al.*¹⁰. This model assumes an exospheric temperature of 1,200 K, a solar flux at line centre at the Earth of $5.14 \times 10^{11} \text{ photons cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1}$ (ref. 11), and no absorption by interplanetary H along the path to and from Saturn. The assumed temperature value is uncertain, but 1,200 K is consistent with the Pioneer 11 radio occultation data¹². A disk-averaged $\text{Ly}\alpha$ intensity of 700–800 R implies that a homopause value of the eddy diffusion coefficient is in excess of $10^7 \text{ cm}^2 \text{ s}^{-1}$. Uncertainties about the role of particle precipitation and the methane density distribution and temperature structure of the mesosphere and thermosphere prevent an exact calculation of the eddy diffusion coefficient at this time. Interplanetary absorption of $\text{Ly}\alpha$ on the round trip to Saturn has been modelled by Bertaux *et al.*¹³, who predict 40% absorption at the time of these observations; thus, Saturn's $\text{Ly}\alpha$ emission could be significantly brighter as measured by a spacecraft in the vicinity of the planet. An H column of 10^{14} cm^{-2} is optically thick and Saturn should exhibit limb darkening at $\text{Ly}\alpha$; however, with the existing spatial resolution it is not possible to distinguish between a cosine darkening function and a uniformly emitting disk. Note that the points plotted in Fig. 1 represent average brightness over an area extending 10 arc s in the east-west direction; assuming a cosine darkening east-west (as well as north-south), the actual sub-solar point brightness would be ~5% higher than that measured.

Exposures were also taken of Titan near western elongation on 5 May 1980, of the sky background 4 arc min north of Saturn on 9 May 1980, and of the region of the outer rings 20 arc s east of the centre of Saturn on 9 May 1980. To within ± 200 R no $\text{Ly}\alpha$ emission above the level of sky background was detected from either Titan or the rings, although the ring exposure was taken on a day when no bursts of $\text{Ly}\alpha$ emission above the 700–800 R level were detected on the saturnian disk. The additional sky background exposure was taken to check for a possible extended H atmosphere around the saturnian system, as hypothesized by McDonough and Brice⁶; however, the background brightness 4 arc min north of Saturn was the same within ± 200 R as that measured immediately north and south of Saturn in the preceding and following exposures. All exposures to date are listed in Table 1, specifying the target area, the longitude of the central meridian of Saturn, the tilt angle of the ring plane north or south of the line of sight from Earth, and the observed level of emission above the 700–800 R planetary level. The longitudes listed are in the Saturn longitude system (λ_{SLs}) defined by the recent observations by the Voyager Planetary Radio Astronomy Group¹⁴. A comparison was also made between the data in Table 1 and the occurrence of radio emission from Saturn, but no direct correlation was discovered.

Unfortunately, the angular resolution of the IUE is not sufficient to pin down the source of the enhanced $\text{Ly}\alpha$ emission.

The excess emission from the southern half of the saturnian disk observed on 19 January, for example, could have originated anywhere along the shaded region shown in Fig. 2 and satisfied our viewing constraints. However, because the rings were within $\pm 2^\circ$ of being edge-on to the Sun throughout these observations and the $\text{Ly}\alpha$ bursts were variable on a time scale of days, it seems unlikely that photo-sputtering of ring particle ice is a plausible source mechanism for these bursts. In addition, <200 R of $\text{Ly}\alpha$ emission was seen from the vicinity of Titan on the same day that a burst of brightness 600 R was detected from the northern half of the saturnian disk, implying that the source of this enhanced emission was not Titan or an associated torus. The most plausible production mechanisms seem to be either: (1) aurora in Saturn's upper atmosphere or (2) a cloud of H atoms in the vicinity of the rings which are released by charged particle sputtering and radiate through resonant scattering of solar $\text{Ly}\alpha$ radiation (and possibly through charged particle excitation). More accurate knowledge of the source of these bursts of $\text{Ly}\alpha$ emission requires better spatial resolution and more flexibility in the angle of view, as will be available from the Voyager 1 and 2 spacecraft now approaching encounter with Saturn.

Table 1 Observations of H Ly emission from Saturn

Date (1980)	Target	Ring plane tilt angle	λ_{SLs} (CML)	Average background level (kR)	Emission above model level (R)
19 January	5° S	1.6° S	319°	1.5	+900
	5° S	1.6° S	346°	1.5	+900
	5° N	1.6° S	34°	1.5	—
12 March	8° N	0°	306°	2.5	—
	Centre	0°	338°	2.5	—
13 March	8° S	0°	10°	2.5	—
3 May	8° N	1.5° N	315°	1.4	+300
	8° S	1.5° N	350°	1.4	+300
	Centre	1.5° N	25°	1.4	+300
5 May	8° N	1.5° N	355°	3.0	+600
	Centre	1.5° N	32°	3.0	—
	8° S	1.5° N	75°	3.0	—
	Titan	1.5° N	117°	3.0	—
9 May	8° N	1.5° N	76°	1.8	—
	4° N	1.5° N	112°	1.8	—
	8° S	1.5° N	149°	1.8	—
	20° E	1.5° N	189°	1.8	—
22 July	7° N	0°	118°	2.0	—
	7° S	0°	155°	2.0	—
	7° N	0°	189°	2.0	—
	Centre	0°	223°	2.0	—
	7° N	0°	262°	2.0	—
	7° N	0°	296°	2.0	—

We thank the IUE Observatory staff for acquisition and reduction of the satellite data. This research was supported by NASA under grant NSF 5393 to the Johns Hopkins University. S.K.A. acknowledges support received from a NASA-Planetary Atmospheres Grant to the University of Michigan.

Received 16 October; accepted 4 December 1980.

- Weiser, H., Vitz, R. C. & Moos, H. W. *Science* **197**, 755–757 (1977).
- Barker, E. S., Cazes, S., Emerich, C., Vidal-Madjar, A. & Owen, T. *Astrophys. J.* **242**, 383–394 (1980).
- Judge, D. L., Wu, F.-M. & Carlson, R. W. *Science* **207**, 431–434 (1980).
- Cheng, A. F. & Lanzerotti, L. J. *J. geophys. Res.* **83**, 2597–2602 (1978).
- Carlson, R. W. *Nature* **283**, 461 (1980).
- McDonough, T. R. & Brice, N. M. *Icarus* **20**, 136–145 (1973).
- Bogges, A. *et al. Nature* **275**, 377–385 (1978).
- Lane, A. L. *et al. Nature* **275**, 414–415 (1978).
- Bertaux, J. L., Moos, H. W., Atreya, S. K. & Lane, A. L. *Astrophys. J. Lett.* **241**, 179–182 (1980).
- Clarke, J. T. *et al. Astrophys. J.* **249**, 696–701 (1980).
- Mount, G. H., Rottman, G. J. & Timothy, J. G. *J. geophys. Res.* **85**, 4271–4274 (1980).
- Kliore, A. J. *et al. Science* **207**, 446–449 (1980).
- Bertaux, J. L., Festou, M., Barker, E. S. & Jenkins, E. B. *Astrophys. J.* **238**, 1152–1159 (1980).
- Desch, M. D. & Kaiser, M. L. (in preparation).
- Luhmann, J. G. & Walker, R. J. *J. geophys. Res.* (in the press).

Mechanical aggregation of enstatite chondrites from an inhomogeneous debris cloud

C. A. Leitch & J. V. Smith

Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA

Enstatite chondrites have oxygen isotope ratios closer to those of the Earth and Moon than other meteorites. Their minerals are chemically reduced; metal contains Si, and some Ca, Ti, Mg and Mn are incorporated in sulphides rather than silicates. Clinostatite and olivine are virtually Fe-free^{1,2}. We have reported³ two types of clinostatite in the Indarch enstatite chondrite, one luminescing blue and one red. Similar clinostatites in the Kota-Kota enstatite chondrite are associated with two distinct types of forsteritic olivine, one luminescing orange and the other blue. The textural relations and differences in chemical composition cannot be explained by progressive condensation from the solar nebula and require the mechanical mixing of material from at least two sources. We suggest here that enstatite chondrites result from mechanical and chemical processes during aerodynamic sorting and gravitational settling of debris from a hot cloud of dust, liquid and gas produced during collision of planetesimals.

Kota-Kota, like Indarch, contains four types of clinostatite-rich chondrules: porphyritic > fine-grained > barred > radiating. One radiating chondrule composed primarily of SiO₂ was observed. A luminescence microscope revealed distinct red and blue clinostatite grains, as well as orange and blue olivine grains. The latter were not found in the first thin section of the Indarch chondrite³, but a few orange and blue grains were found in six new sections. Porphyritic chondrules are extremely

complex and a single chondrule may contain all four phases (Fig. 1). Red clinostatite grains may enclose blue clinostatite, orange olivine and/or blue olivine, while individual blue clinostatite grains may enclose orange and/or blue olivine. Red clinostatite was not observed inside blue clinostatite. Fine-grained chondrules contain blue clinostatite and rare blue olivine. Both barred and radiating chondrules are composed predominantly of lamellae of either red or blue clinostatite. All chondrule types are surrounded by a shell composed of a fine-grained mixture of all minerals found in the meteorite. Similar shells were found surrounding chondrules in Indarch³ and in ordinary chondrites⁴.

Electron microprobe analyses were obtained with an ARL-EMX instrument. Magnesium, Si and Fe were determined with a solid state detector at 15 kV, 0.1 μ A and 120 s live time. Minor elements were determined with crystal spectrometers and 25 kV and 2 μ A. Counting times were selected for peaks and backgrounds to achieve a precision near 30 p.p.m. (1 σ) for individual analyses. Standards were: Ca₂P₂O₇(P), Di₈₅Jd₁₅(Ca, Al, Na), Corning glass V(Ti, Cr), Corning glass W(Mn) and Corning glass X(Ni).

As for Indarch³, minor element contents correlate with luminescence colour and not with chondrule type. Representative analyses in Table 1 are taken only from the chondrule in Fig. 1. Table 2 shows observed ranges in minor element contents for both olivine and clinostatite. Red clinostatites have

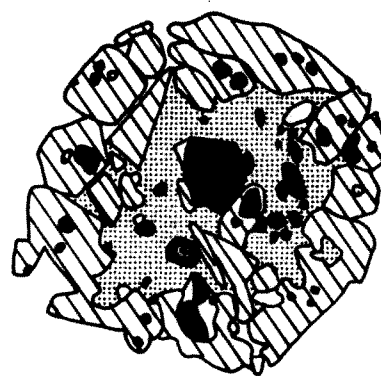


Fig. 1 Schematic representation of a porphyritic chondrule. Black, clear, hatched and stippled areas are respectively, orange olivine, blue olivine, red clinostatite and blue clinostatite. Blue clinostatite is polycrystalline, but individual grain boundaries are not shown. The chondrule is 410- μ m wide.

Table 1 Olivine and pyroxene analyses from one Kota-Kota chondrule

	1	2	3	4	5	6
P ₂ O ₅	0.005	0.004	0.005	0.005	NA	<0.1
SiO ₂	42.3	42.2	58.9	58.9	58.9	58.2
TiO ₂	0.016	0.024	0.079	0.008	NA	0.03
Al ₂ O ₃	0.051	0.077	0.39	0.108	0.27	0.32
Cr ₂ O ₃	0.23	0.064	0.32	0.024	NA	0.36
FeO _T	0.96	0.86	0.95	0.81	0.87	0.96
MnO	0.123	0.033	0.20	0.028	0.13	0.1
MgO	56.2	56.3	38.6	38.9	38.7	39.5
NiO	0.003	0.001	0.005	0.003	NA	0.034
CaO	0.167	0.122	0.28	0.088	0.32	0.4
Na ₂ O	0.0	0.0	0.021	0.107	0.09	0.05
Sum	100.055	99.685	99.750	98.881	100.28	100.0
Cations/24 oxygens						
P	0.0006	0.0005	0.0006	0.0006	—	—
Si	5.9769	5.9769	7.9496	7.9898	7.9762	7.864
Ti	0.0017	0.0026	0.0080	0.0008	—	0.004
Al	0.0085	0.0129	0.0620	0.0173	0.0431	0.052
Cr ³⁺	0.0257	0.0072	0.0341	0.0026	—	0.040
Fe ²⁺	0.1134	0.1019	0.1072	0.0919	0.0985	0.108
Mn	0.0147	0.0040	0.0229	0.0032	0.0149	0.012
Mg	11.8362	11.8853	7.7652	7.8652	7.8114	7.948
Ni	0.0003	0.0001	0.0005	0.0003	—	0.004
Ca	0.0253	0.0185	0.0405	0.0128	0.0464	0.056
Na	0.000	0.000	0.0055	0.0281	0.0236	0.012

Columns 1–6, electron microprobe analyses (iron given as FeO): 1, orange luminescing olivine, mean of four analyses. 2, Blue luminescing olivine, mean of three analyses. 3, Red luminescing clinostatite, mean of seven analyses. 4, Blue luminescing clinostatite, mean of three analyses. 5, Mean of 27 clinostatite analyses from ref. 2. 6, Mean of analyses of clinostatite from ref. 1. 6, Wet chemical analysis of bulk pyroxene separate⁵; FeO 0.93, Fe₂O₃ 0.03 wt %. NA, not applicable.

higher concentrations of Ti, Al, Cr, Mn and Ca, whereas blue ones have higher Na. Orange olivines contain more Cr and Mn than blue ones.

The method of assignment of valence states to minor elements in the pyroxenes has been detailed elsewhere³. There is insufficient Al and Cr³⁺ in blue clinostatite to balance the observed Na, and some Fe must be assigned to the trivalent state to maintain charge balance if a defect-free MTO₃ formula is assumed. For red clinostatites, it was again necessary to assume that some Fe³⁺ is present to obtain a cation: oxygen ratio fitting the theoretical value of 2:3. For clinopyroxene analyses in columns 3 and 4 of Table 1, it is necessary to assume that 38 and 8%, respectively, of total Fe is present as Fe³⁺. However, a wet chemical analysis of a pyroxene separate from Kota-Kota⁵ showed only 0.03 wt % Fe₂O₃ against 0.93 wt % FeO. The trivalent Fe was attributed to terrestrial oxidation, and the significance of the recalculation of the electron microprobe analyses is unclear. If some or all of the Cr were divalent, even more Fe³⁺ would be required.

The observed textures and bimodal chemical compositions of the olivines and clinostatites rule out formation of these meteorites by direct condensation from the solar nebula. Mechanical aggregation of material from at least two source regions is needed to explain the blue and red varieties, and the

Table 2 Observed compositional variations

	Kota-Kota olivine orange	Kota-Kota pyroxene red	Indarch olivine orange	Indarch pyroxene red
P ₂ O ₅	0.003–0.008	0.000–0.010	0.000–0.007	0.000–0.017
TiO ₂	0.000–0.022	0.025–0.120	0.000–0.085	0.018–0.077
Al ₂ O ₃	0.016–0.100	0.151–0.451	0.033–0.119	0.119–0.390
Cr ₂ O ₃	0.115–0.290	0.126–0.568	0.108–0.322	0.117–0.681
MnO	0.090–0.188	0.091–0.418	0.093–0.209	0.086–0.513
NiO	0.000–0.005	0.000–0.010	0.000–0.005	0.000–0.015
CaO	0.090–0.252	0.124–0.472	0.116–0.272	0.123–0.419
Na ₂ O	0.000	0.002–0.093	0.000	0.000–0.114
	blue	blue	blue	blue
P ₂ O ₅	0.002–0.009	0.001–0.009	0.000–0.008	0.000–0.017
TiO ₂	0.002–0.031	0.001–0.012	0.000–0.047	0.000–0.014
Al ₂ O ₃	0.041–0.116	0.015–0.195	0.034–0.141	0.012–0.172
Cr ₂ O ₃	0.029–0.095	0.002–0.107	0.020–0.087	0.005–0.098
MnO	0.014–0.046	0.000–0.090	0.010–0.062	0.001–0.086
NiO	0.001–0.004	0.000–0.013	0.000–0.006	0.000–0.006
CaO	0.082–0.188	0.015–0.157	0.094–0.190	0.015–0.119
Na ₂ O	0.000	0.040–0.248	0.000	0.049–0.250

Electron microprobe analyses, wt %. Accuracy is 2–5% of amount present with detection level near 0.005 (2 σ).

observed textures indicate a complex sequence of events. Enclosure of rounded olivine grains by pyroxene grains suggests that olivine formed earlier than clinoenstatite in agreement with condensation calculations for a solar nebula with either the solar C/O ratio of 0.57 (ref. 6) or the higher ratio of ~1.0 postulated⁷ to explain enstatite chondrites.

Clinoenstatite melts incongruently at 1,557 °C to form forsterite plus liquid⁸. Because the forsterite bodies are rounded rather than euhedral, it is unlikely that the forsterite observed in Kota-Kota was so formed. In addition, both orange and blue olivine occur surrounded by red or blue pyroxene. Because forsterite melts at a much higher temperature (1,850 °C) than enstatite, pre-existing forsterite might have survived the melting of clinoenstatite.

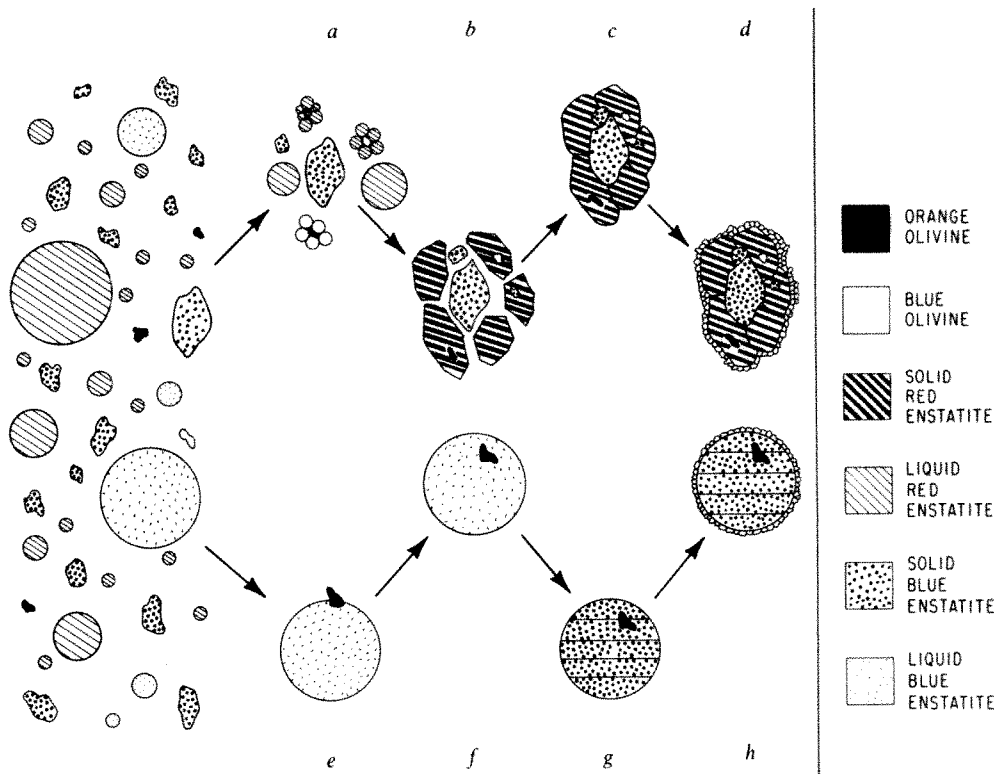
The simplest scheme to be considered for the Kota-Kota meteorite involves (1) progressive condensation of two separate

regions of the solar nebula to give minor forsterite and dominant clinoenstatite, and (2) mechanical mixing of the condensates. Because this scheme is insufficient to explain the textural properties of the chondrules, some other process or processes must be involved. In view of evidence^{9–11} for accretion of planets through planetesimals, generation of chondrules during collision of planetesimals is considered¹².

The similar ranges of clinoenstatite compositions in Table 2 indicate that the Indarch and Kota-Kota meteorites may have come from the same parent body. Orange olivine and red clinoenstatite are assumed to derive from the same condensation region because of comparable Cr₂O₃ and MnO contents, and similarly the blue olivine and blue clinoenstatite are paired. Enclosure by red clinoenstatite of blue clinoenstatite, orange olivine and blue olivine rules out mechanical mixing of crystals that condensed in equilibrium conditions in two source regions before planetesimal collision. To explain the textural relations it seems that at least some red clinoenstatite must have crystallized after the other three minerals. Because there is a sharp discontinuity in luminescence and chemistry between adjacent grains of porphyritic chondrules (Fig. 1), it seems that an enclosing grain grew without reaction with an enclosed grain. Figure 2a–d shows how the porphyritic chondrules may have formed, partly by mechanical aggregation of pre-existing crystals and partly by growth of red clinoenstatite around pre-existing crystals. A liquid stage for at least some red grains seems necessary to explain the inclusion of blue clinoenstatite and orange and blue olivines, and it may prove necessary to introduce additional steps of crystallization and mechanical disaggregation. Interstices between aggregated crystals are assumed to be filled by dust and vapour deposits.

Barred, radiating and fine-grained chondrules are assumed to have derived from liquids formed during planetesimal collision. Because these droplet chondrules are composed predominantly of either red or blue clinoenstatite with only minor amounts of interstitial phases, it is necessary to postulate formation of two separate liquids with compositions near those of the red or blue clinoenstatites. Such liquid droplets could then incorporate minor amounts of pre-existing solid material released by mechanical disaggregation of the colliding planetesimals (Fig. 2e–h). The impact melts are assumed to have been dispersed by high velocity jets¹³ produced during collision. Such jets must

Fig. 2 Schematic representation of chondrule formation from a gas-dust-liquid droplet cloud. Angular shapes are crystals; circles are liquid. a–d shows the formation of a porphyritic chondrule: a, some blue enstatite, blue olivine and orange olivine becomes surrounded by red liquid; b, red liquid forms crystals enclosing blue enstatite, blue olivine and orange olivine; c, crystals become aggregated and a porphyritic chondrule similar to that in Fig. 1 is formed; d, chondrule becomes surrounded by a shell of fine-grained dust. e–h shows the formation of a barred chondrule with an olivine inclusion: e, blue liquid droplet comes in contact with an orange olivine; f, olivine is trapped inside blue liquid droplet; g, liquid crystallizes and assumes a typical barred appearance; h, chondrule becomes surrounded by a shell of fine-grained material.



have dispersed the molten material quickly enough to prohibit mixing between the two liquids.

In such a scheme, the massive debris from the collision is assumed to coalesce into a new planetesimal in part of which melting and high-grade metamorphism would occur¹². Smaller fragments would remain in orbit, and later fall onto the surface. Dodd¹⁴ has given evidence for aerodynamic sorting of chondrules in ordinary chondrites⁴. He argued in favour of continuous entry of particles into an atmosphere because gas drag in a stagnant atmosphere would produce a graded deposit on the surface of a growing planetesimal. However, shock waves produced in a gas-dust cloud by incoming bodies from colliding planetesimals may have caused material to settle rapidly into an ungraded deposit¹².

Although details still need to be worked out, it seems certain

that enstatite chondrites are not direct products of the solar nebula. In spite of the complexity of the mineralogy and petrography, we need not abandon the well-documented mixing models for the bulk compositions of unequilibrated enstatite^{15,16} chondrites, at least as a first approximation. Provided that little material was lost to the solar nebula during collision, the bulk chemistry of chondrites should be explicable in terms of direct condensates that entered the original planetesimals. Although the present scheme applies specifically to enstatite chondrites, the general features may also apply to the unequilibrated ordinary chondrites^{17,18}.

We thank R. T. Dodd for helpful criticism, E. Olsen and K. Keil for specimens, NASA for grants NSG-7550 and 14-001-171, and I. M. Steele, R. Draus, I. Baltuska and O. Draughn for technical assistance.

Received 30 July 1980; accepted 8 January 1981.

1. Keil, K. *J. geophys. Res.* **73**, 6945 (1968).
2. Mason, B. *Geochim. cosmochim. Acta* **30**, 23 (1966).
3. Leitch, C. A. & Smith, J. V. *Nature* **283**, 60 (1980).
4. Allen, J. S., Nozette, S. & Wilkening, L. L. *Lunar planet. Sci.* **10**, 27 (1979).
5. Binns, R. A. *Miner. Mag.* **37**, 649 (1970).
6. Grossman, L. & Larimer, J. W. *Rev. Geophys. Space Phys.* **12**, 71 (1974).
7. Larimer, J. W. & Bartholomay, M. *Geochim. cosmochim. Acta* **43**, 1455 (1979).
8. Bowen, N. L. & Andersen, O. *Am. J. Sci.* 4th Ser. **37**, 488 (1914).

9. Safronov, V. S. *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets* (Israel Program for Scientific Translations, Jerusalem, 1972).
10. Wetherill, G. W. *Proc. 7th Lunar Sci. Conf.* 3245 (1976).
11. Smith, J. V. *Miner. Mag.* **43**, 1 (1979).
12. Smith, J. V. *Lunar planet. Sci.* (in the press).
13. Kieffer, S. *Science* **189**, 333 (1975).
14. Dodd, R. T. *Earth planet. Sci. Lett.* **30**, 281 (1976).
15. Larimer, J. W. & Anders, E. *Geochim. cosmochim. Acta* **31**, 1239 (1967).
16. Ikramuddin, M., Binns, C. M. & Lipschutz, M. E. *Geochim. cosmochim. Acta* **40**, 133 (1976).
17. Hutchison, R., Bevan, A. W. R., Agrell, S. O. & Ashworth, J. R. *Nature* **280**, 116 (1979).
18. Bevan, A. W. R. & Axon, H. J. *Earth planet. Sci. Lett.* **47**, 353 (1980).

Porous anodic film formation on aluminium

G. E. Thompson & G. C. Wood

Corrosion and Protection Centre, University of Manchester Institute of Science and Technology, Manchester M60 1QD, UK

The development of porous anodic films on aluminium in the major forming acids has been established¹⁻⁵. However, little information has been provided to explain the markedly different rates of steady-state film growth for anodizing in constant voltage conditions in these acids, with the anodic film cell parameters dependent on the cell voltage¹. The regular cellular morphology of the steady-state films and the substructure developed in the forming acids are revealed in Fig. 1. The characteristic parallel-sided pores, at the centres of approximately hexagonal cells, pass normal to the macroscopic metal surface but are separated from it by the barrier layer. Here, from considerations of the film material substructure and the distribution of species within the film derived from the forming acid, it is postulated that the likely variations in the local field strengths across the barrier layers can explain the steady-state anodizing behaviour observed in the different acids in constant voltage conditions.

Direct transmission electron microscopy of ion-beam thinned steady-state porous anodic films, formed originally in phosphoric or oxalic acid, has shown the presence of cell boundary bands separated from the electrolyte in the pores by differently textured film material⁶ (Fig. 1). Microanalysis of the regions comprising the cell material of the film formed in phosphoric acid suggests that the cell-boundary band is composed of virtually pure alumina, whereas phosphorus is detected in the adjacent differently textured material⁷. For films formed in chromic acid, no significant chromium is detected in the cell material⁵, which appears similar in nature to the cell boundary band material observed for films formed in phosphoric acid, that is, relatively pure alumina. The highly acid anion contaminated films formed in sulphuric acid also show a region of relatively pure alumina which is not resolved readily in the electron microscope for the films formed at the usual low voltages (R. Ozcan and our unpublished work). From such knowledge it is possible to depict the barrier layers of the films formed in the major anodizing acids in terms of the ratio of the thicknesses of

relatively pure alumina region to acid anion-contaminated material, which seems to increase in the order: sulphuric acid (0.05) > oxalic acid (0.1) > phosphoric acid (0.5) > chromic acid (∞).

The relative extents of the different regions present in the barrier layers of the films formed at constant voltages in the usual anodizing conditions in the above acids are illustrated in Fig. 1, which shows that the films form at relatively similar nm V⁻¹ ratios. The relatively pure alumina is drawn as a solid region as it is thought that this more compact film grows by ionic migration to form a glassy or anion-free microcrystalline structure, whereas the outer microcrystalline acid anion-contaminated film material is deposited under the field from initially colloidal alumina⁸. Different extents of deposition explain the varying current efficiencies observed in the corresponding acids.

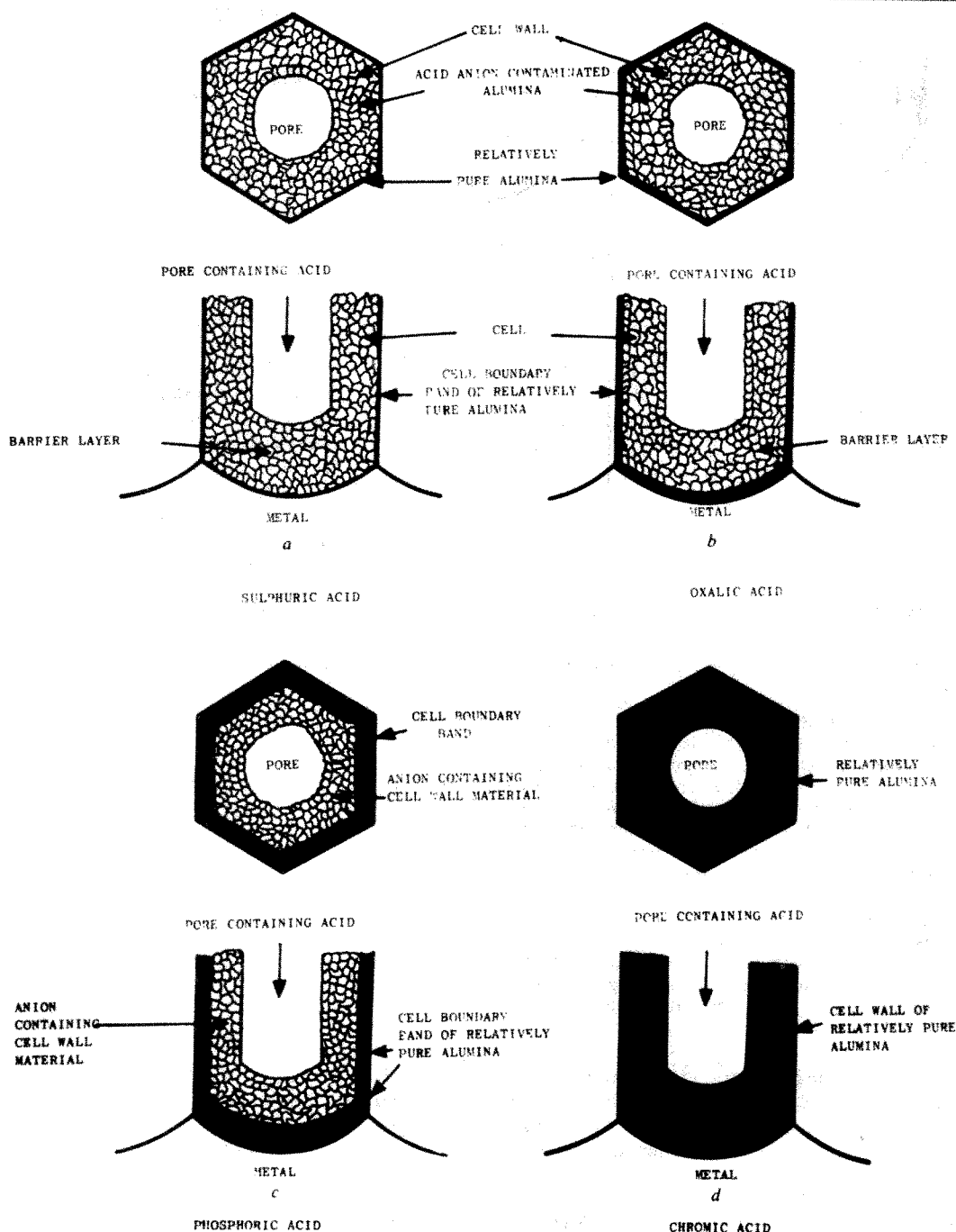
There is a direct link between the above findings and the rates of formation of the films in the different anodizing acids, at constant voltage. For films formed at 20 V, typical rates of formation increase in the order: sulphuric acid > oxalic acid > phosphoric acid > chromic acid.

For example, the rate of formation of the film in sulphuric acid, with the thinnest relatively pure alumina region, is much greater than in chromic acid, where films composed of relatively pure alumina are formed. Furthermore, the current density (i) is related to the potential drop (V) across the barrier layer by the high field conduction theory:

$$i = A \exp (BV/d)$$

where A and B are constants, for a given temperature, and d is the overall barrier layer thickness. In the above equation V/d represents the effective average field across the barrier layer which is approximately constant for the barrier layers of the different films because the measured nm V⁻¹ ratios are similar. However, if the true field across the relatively pure alumina region of the barrier layer is in the order: sulphuric acid > oxalic acid > phosphoric acid > chromic acid, because this represents the order of increasing thickness of the observed cell boundary bands, then the rates of formation of the films in the different acids will also vary in the order given above if solid state ionic migration across the inner layer is the rate-determining step. The passage of ions through the outer regions of the contaminated film is thought to proceed relatively easily. Thus, during anodizing in the major anodizing acids, the field strength is thought to be greater across the inner, compact relatively pure

Fig. 1 The plan and sectional views of a pore and adjacent cell for films formed in each of the major acids: *a*, sulphuric acid; *b*, oxalic acid; *c*, phosphoric acid; *d*, chromic acid. The cell comprises acid anion-contaminated film material adjacent to the pore and relatively pure alumina remote from it: □ acid anion contaminated film material; ■, relatively pure alumina.



alumina region and lower across the acid anion-contaminated alumina film, which probably varies from solid in contact with the cell boundary bands to gel-like material at the base of the pores. Microporosity or relatively easy passage of ions through intercrystallite pathways will reduce the effective field further. Figure 2 shows how the total voltage drop across the barrier layer is similar for the highly contaminated film formed in sulphuric acid (Fig. 2a) and the acid anion-free film formed in chromic acid (Fig. 2d) where the cell boundary band to contaminated film material ratios approach 0 and ∞ , respectively. The current density is higher in the former case because of the higher voltage drop across the pure alumina region adjacent to the metal interface. In addition, anodizing in chromic acid proceeds at relatively low current efficiency. In phosphoric and oxalic acids there are two zones. The voltage drop and field across the pure cell boundary bands are high but low in the outer acid anion-contaminated film material (Fig. 2b, c).

The variation in voltage drop across the adjacent film regions comprising the barrier layer is important and has been considered by Dell'Oca and Young⁹ in an investigation of the

different layered barrier film formed on tantalum in phosphoric acid. They suggested that a high field operates in the contaminated layer adjacent to the solution where phosphate reduces the permittivity and decreases ionic mobility. The case for a dependence of field strength on film thickness (for thin films) has been put forward by Dignam¹⁰. Here the different field strengths have been based on the likely mechanisms of film growth during steady-state porous film formation involving solid-state ionic migration across the relatively pure alumina region and deposition of the outer layer from a hydrous metal oxide solution. Direct observations of the film materials, revealed by ultramicrotomy of both barrier and porous-type anodic films, show the sections to be relatively uniform and featureless^{2,3,8,9,11}, only revealing telling changes after exposure to the electron beam¹⁰. This is thought to represent 'drying out' and/or ageing of the film material after anodizing, and it is only during anodizing that the variation in field strengths across the barrier layer is manifested. As the nm V^{-1} ratios for the different films have been determined after anodizing in the absence of the field it is not surprising that they show relatively similar values.

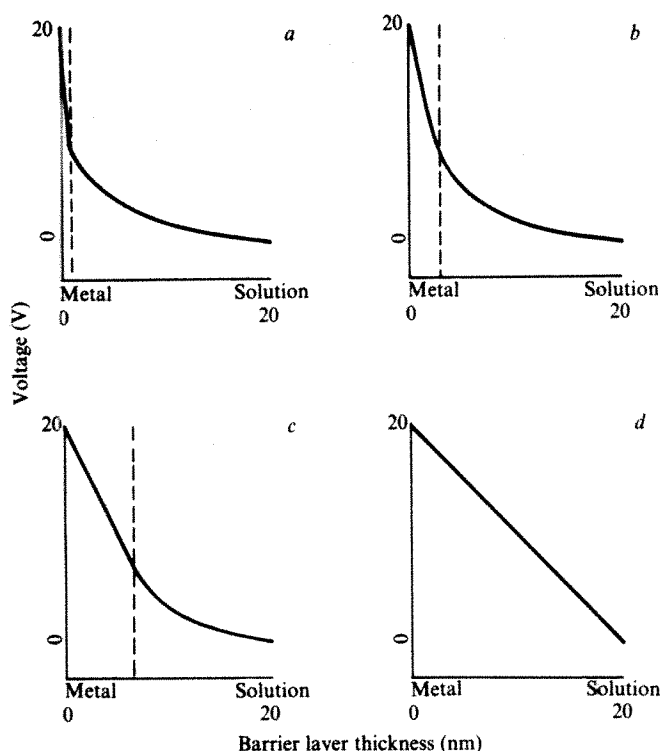


Fig. 2 Distribution of the voltage drop and field (the slope of the voltage-distance plot) across typical barrier layers of porous anodic films formed in each of the major acids: *a*, sulphuric acid; *b*, oxalic acid; *c*, phosphoric acid; *d*, chromic acid. The voltage drop is greatest and linear across the inner compact relatively pure alumina region and lower across the outer acid anion contaminated film material, where the voltage decreases progressively towards the film-solution interface. Consequently, the highest field is across the least extensive relatively pure alumina region, evident for the films formed in sulphuric acid, decreasing for the films formed in oxalic and phosphoric acids respectively, and lowest for the most extensive pure alumina region present in the films formed in chromic acid. The rates of formation of the corresponding films decreases in the same order.

The nature of the incorporated acid anion will probably strongly influence the size distribution of the deposited and contaminated microcrystalline alumina¹², and its hydrogen bonding characteristics, which affect the effective compaction of this material.

Thus, the steady-state anodizing behaviour of porous anodic films formed in the major anodizing acids can be related to the distribution of the acid anions within the barrier layers and the true field strengths across the relatively pure alumina regions. The effect of the acid anion on thermally enhanced field-assisted dissolution, or other processes whereby the relatively pure alumina region transforms to the acid anion-incorporated material and its influence on anodizing at constant current density, will be discussed elsewhere.

Received 18 November 1980; accepted 9 January 1981.

- O'Sullivan, J. P. & Wood, G. C. *Proc. R. Soc. A* **317**, 511-543 (1979).
- Thompson, G. E., Furneaux, R. C., Goode, J. S. & Wood, G. C. *Trans. Inst. Metal Finishing* **56**, 159-167 (1978).
- Thompson, G. E., Furneaux, R. C. & Wood, G. C. *Trans. Inst. Metal Finishing* **55**, 117-128 (1977).
- Bailey, G. & Wood, G. C. *Trans. Inst. Metal Finishing* **52**, 187-199 (1975).
- Thompson, G. E., Wood, G. C. & Hutchings, R. *Trans. Inst. Metal Finishing* **58**, 21-25 (1980).
- Thompson, G. E., Furneaux, R. C. & Wood, G. C. *Corros. Sci.* **18**, 481-498 (1978).
- Thompson, G. E., Furneaux, R. C., Wood, G. C. & Hutchings, R. *J. electrochem. Soc.* **125**, 1480-1482 (1978).
- Thompson, G. E., Furneaux, R. C., Wood, G. C., Richardson, J. A. & Goode, J. S. *Nature* **272**, 433-435 (1978).
- Dell'Oca, C. J. & Young, L. *J. electrochem. Soc.* **117**, 1548-1551 (1970).
- Dignam, M. J. *J. electrochem. Soc.* **126**, 2188-2195 (1979).
- Furneaux, R. C., Thompson, G. E. & Wood, G. C. *Corros. Sci.* **18**, 853-881 (1978).
- Matyevic, E., Bell, A., Brace, R. & McFadyen, P. *J. electrochem. Soc.* **120**, 893-899 (1973).

Tasmanian tree rings belie suggested anthropogenic $^{13}\text{C}/^{12}\text{C}$ trends

R. J. Francey

CSIRO Division of Atmospheric Physics, Aspendale, Victoria, Australia 3195

Published measurements in tree rings¹⁻¹⁰ suggest a persistent decrease with time of the ratio $^{13}\text{C}/^{12}\text{C}$, particularly in the period 1860-1940, and the release of C^{13} -deficient CO_2 from both fossil fuel combustion and forest clearing have often been invoked to explain the decreases. Freyer¹¹ concludes that "the existing measurements leave no doubt that the ^{13}C content in the cellulose fraction of modern wood was indeed decreased significantly". Data presented here for two tree species from sites with contrasting ecologies on the remote and exposed Tasmanian west coast do not, however, support claims that these trends, which have been observed in (mostly) Northern Hemisphere trees, are the global response of the atmosphere to agricultural and industrial activities.

Freyer's statement requires qualification. There is evidence^{2,4,7-9} of a levelling or an increase in $^{13}\text{C}/^{12}\text{C}$ ratios after 1940, moreover the period before 1860 is poorly sampled, thus the statistical significance of a modern trend over several decades is difficult to assess. The Freyer's conclusion is based

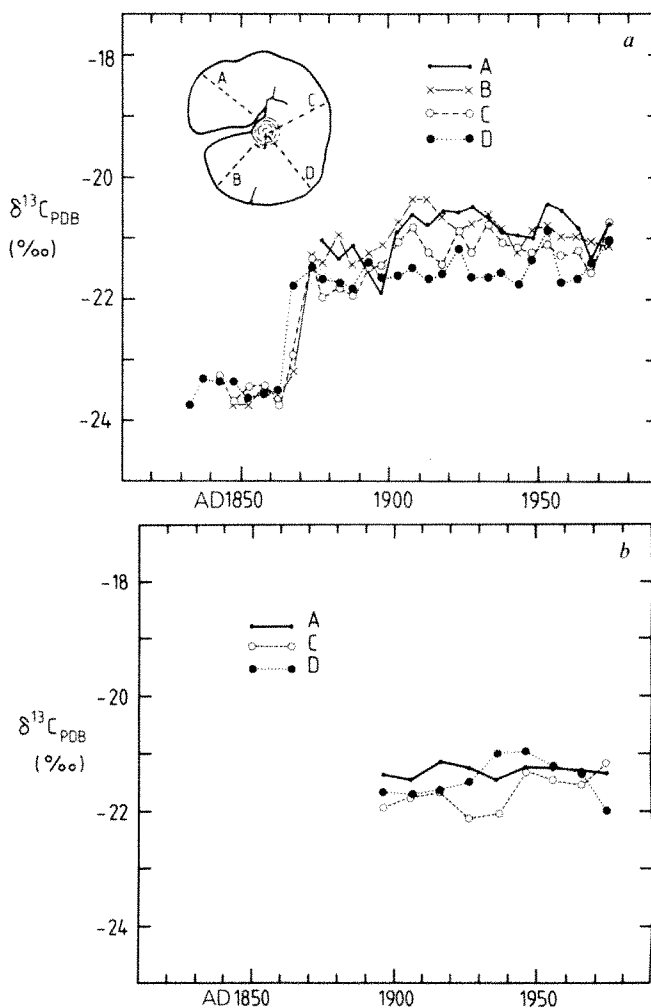


Fig. 1 *a*, $^{13}\text{C}/^{12}\text{C}$ analyses on 5-yr blocks from four radii of tree JR39. Radii were cut from a disk as shown in the inset. *b*, $^{13}\text{C}/^{12}\text{C}$ analyses on 10-yr blocks from three radii of tree BH858.

Table 1 Tasmanian trees analysed for $\delta^{13}\text{C}$

Tree	Species	Site description		Ecology	Approximated germination (AD)	Period analysed (AD)
		Altitude (m)	Distance to ocean* (km)			
SA†	<i>Athrotaxis selaginoides</i>	880	90	Rain forest	<1700	1841–1974
S3†	<i>Athrotaxis selaginoides</i>	880	90	Rain forest	<1700	1893–1974
SC15‡	<i>Athrotaxis selaginoides</i>	880	90	Rain forest	1090	1540–1956
SC19‡	<i>Athrotaxis selaginoides</i>	880	90	Rain forest	1677	1743–1952
MR3.4‡	<i>Athrotaxis selaginoides</i>	1,100	30	Exposed mountain top	<1450	1500–1975
BH858§	<i>Phyllocladus aspleniifolius</i>	40	15	Small island	1880	1885–1958
JR39§	<i>Phyllocladus aspleniifolius</i>	60	35	Rain forest	1829	1830–1977

* In a southwesterly direction.

Dendrochronology from: † D. Adamson (Macquarie University); ‡ J. Ogden (University of Auckland, formerly Australian National University);

§ T. Bird (CSIRO Division of Forestry).

largely on trees of mid-latitude Northern Hemisphere origin, close (in a global sense) to major anthropogenic sources of CO_2 and associated pollutants¹², embedded in large natural seasonal variations in atmospheric CO_2 , and remote from major exchange reservoirs in the southern oceans.

The further step of relating a tree ring derived $^{13}\text{C}/^{12}\text{C}$ trend to the global CO_2 budget presents its own problems⁹, including (1) ignorance of the physical factors influencing $^{13}\text{C}/^{12}\text{C}$ in cellulose, sometimes creating large differences in cellulose of similar age from different locations within a tree, (2) the influence of local environmental factors, and (3) the relative insensitivity of $^{13}\text{C}/^{12}\text{C}$ in the atmosphere–biosphere–ocean system to gradual changes in the biosphere.

The forests of the Tasmanian west coast offer many advantages in the search for a global perspective on the carbon dioxide history of the atmosphere. The area is predominantly wilderness, remote from large scale industry and exposed to the almost circumpolar ocean fetches of the 'roaring forties'. Seasonal variations in the atmospheric CO_2 concentration are less than a tenth of those at high northern latitudes¹³. The native flora includes several species suitable for ring-width dating^{14–16} and living trees of great age, in stable ecosystems, are available.

Results are now available from the first seven trees analysed, whole wood analyses from three of which have been reported elsewhere^{17,18} and the temperature correlations reported there, while still evident at reduced magnitude, do not conflict with the cellulose results reported here. In what follows, comparisons will be made with measurements on cellulose^{1,3,4,8,9} (though similar conclusions would come from consideration of whole wood analyses^{2,4,7}).

The trees are referred to by their collector's designation, and relevant properties are summarized in Table 1. Full details of the trees and the analyses techniques will be given elsewhere¹⁹. The analysis repeatability on a homogeneous wood sample is typically 0.06‰.

On two of the trees, JR39 and BH858, comparisons have been made of $^{13}\text{C}/^{12}\text{C}$ variations along different radii. Analyses on 5-yr blocks of four radii for JR39 are shown in Fig. 1a, the inset shows the radii cut from a disk.

This record shows sharp discontinuity in $\delta^{13}\text{C}_{\text{PDB}}$ occurring at around 1869, where scarring suggests fire damage. After 1869 there is a corresponding sharp increase in ring width. The $\delta^{13}\text{C}_{\text{PDB}}$ and ring width increases are probably the result of sudden exposure of a sub-canopy sapling to free atmosphere, removing limiting factors related to competition, and/or markedly reducing the uptake of ^{13}C depleted respired carbon dioxide²⁰. It suggests that variations in $\delta^{13}\text{C}$ after this event were not seriously influenced by these factors.

While $\delta^{13}\text{C}_{\text{PDB}}$ varies by up to 1‰ in different radii, these differences are generally maintained throughout the adult life of the tree. In the context of possible large monotonic trends¹¹, this is taken as evidence (with some caution⁹) that trends representative of the whole disk can be inferred from one radius.

This application when only one radius is available is supported by Fig. 1b, showing three cores from tree BH858. The data are analysed in 10-yr blocks. Radius A is a core from a south-west direction, 1.09 m from the base; C is a north-east core from 0.58 m and D a south-east core from 0.81 m. Although a relatively young rainforest tree, there is no juvenile effect. It may be relevant that it is located on a small (~100 m diameter)

Fig. 2 $^{13}\text{C}/^{12}\text{C}$ analyses on all seven trees. Averages of the different radii are used for JR39, BH858. (Dashed lines indicate values corresponding to periods of marked ring width discrepancies.)

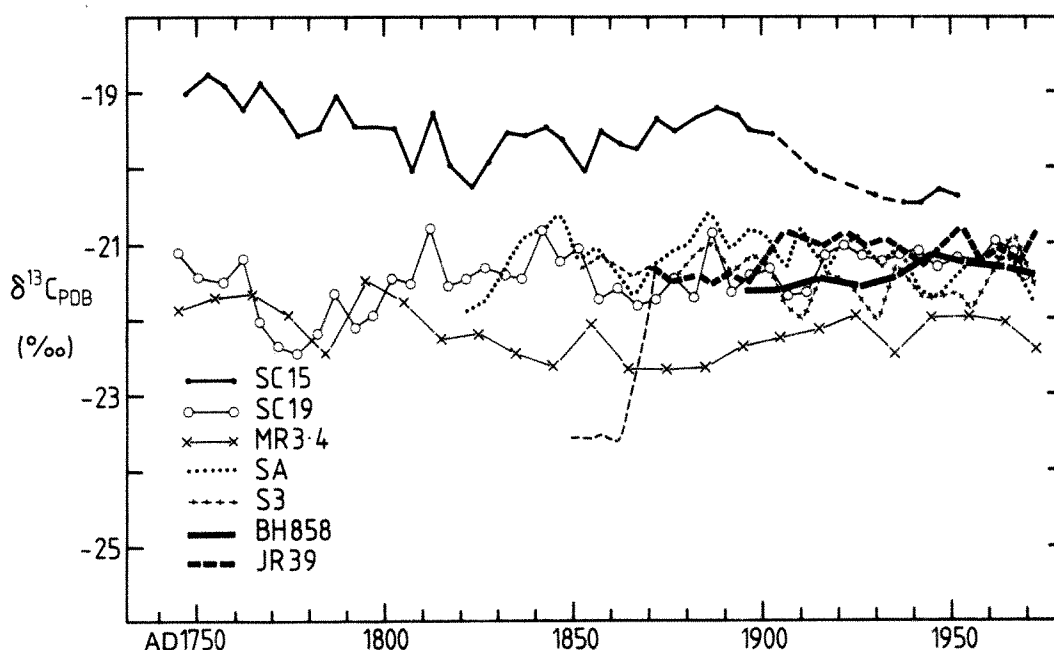


Table 2 Comparison of $\delta^{13}\text{C}$ trends since 1950

Medium	Period of data	Regression coefficient (% yr ⁻¹)	Ref.
Atmosphere	1956–1978	-0.030	21, 22
Cellulose	1950–1972	-0.018 ± 0.008	3
Cellulose	1950–1972	+0.042 ± 0.011	9
Cellulose	1951–1974	-0.005 ± 0.008	This work (6 tree averages)

exposed island in Bathurst Harbour (separated by 15 km of water and grassed hills from the open ocean), and flushing of respired sub-canopy CO_2 should be much more efficient than in a land-locked, heavily vegetated area.

Figure 2 shows an average curve for each of these trees, the remaining five curves referring to a single radius per tree. Four of these trees (SC15, SC19, SA, S3) are from a rainforest on a mountain ridge well exposed to the south-west, but 90 km from the coast. Three had germination dates estimated at about 1700 (although samples relating to a briefer period were provided for analysis) and were probably well exposed to free atmosphere from 1800. SC15, however, is much older with an estimated germination about 1090. Ring width data indicate a severe setback in growth from 1907 to 1917. In blocks after 1907, $\delta^{13}\text{C}_{\text{PDB}}$ values are on average about 1‰ more negative than immediately before 1907. At the time of sampling SC15 was lying on the ground and the tree may have fallen (had achieved a sub-canopy status) as early as 1907.

There is no possibility of canopy influences over the past several hundred years for MR3.4: estimated germination precedes 1450, and the high altitude, stunted 'krumholtz' form is evidence of its isolation and extreme exposure.

Figure 3 shows the combined data although the first 30 yr of JR39 analyses and the last 50 yr of SC15, (indicated by dashed lines in Fig. 2) for which independent ring width evidence suggest sub-canopy influences, are excluded. SC15, SC19, MR3.4 have common records between 1740 and 1905; and in Fig. 3a averages (in 10-yr blocks) of their deviations from the 1740–1900 mean are shown.

Six trees have common records between 1890 and 1970, and five of these extend back to 1870. Figure 3b shows averages (in 10-yr blocks) of their deviations from the 1890–1970 means (using five trees before 1890 and six trees thereafter).

In the context of global changes in atmospheric $\delta^{13}\text{C}$, the average of the six trees in Fig. 3b exhibits a striking lack of systematic behaviour—the data for a 10-yr interval never deviate significantly from the 1870–1970 mean. This contrasts with

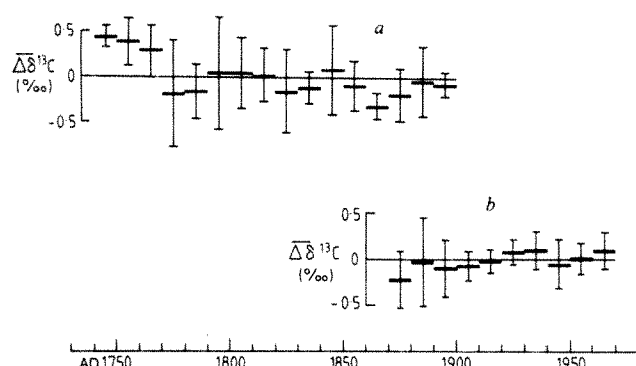


Fig. 3 Combined $^{13}\text{C}/^{12}\text{C}$ data from 10-yr blocks. *a*, The average and standard deviation of deviations from the 1740 to 1900 mean for SC15, SC19, MR3.4. *b*, The average and standard deviation of deviations from the 1890–1970 means for SC19, MR3.4, JR39, BH858, SA and S3; for the period 1870–1890 only BH858 is not represented.

the steady decrease of about 0.5‰ (from 1870 to 1960) of Freyer¹¹.

A large decrease in $\delta^{13}\text{C}$ of almost 1‰ between 1840 and 1920 of Freyer¹¹, frequently interpreted as a global atmospheric response to a large net transfer of CO_2 from the biosphere^{5,6,8}, is not supported by the results in Fig. 3a. Clearly Fig. 3 and Freyer's results, for example, cannot both represent global atmospheric $^{13}\text{C}/^{12}\text{C}$ behaviour.

All Northern Hemisphere cellulose data are not incompatible with the Tasmanian results; over the same 1870–1970 period of Fig. 3b, the mean curve of Tans and Mook⁹ (and the whole wood data of Farmer⁷ and Galimov⁴) show a similar absence of a monotonic trend.

The data of Keeling *et al.*^{21,22} provide a useful comparison with the above results. They report annual average values of $\delta^{13}\text{C}$ and μ (the total CO_2 concentration);

$$1956 \mu = 314.1 \text{ p.p.m.} \quad \delta^{13}\text{C} = -6.69\text{‰}$$

$$1978 \mu = 334.2 \text{ p.p.m.} \quad \delta^{13}\text{C} = -7.34\text{‰}$$

representing a $\delta^{13}\text{C}$ change of -0.6‰ over 22 yr, or -0.030‰ yr⁻¹ for a linear trend.

Table 2 shows the available individual measurements of $\delta^{13}\text{C}$ in cellulose from 1950 linearly regressed against time. The anticipated atmospheric $\delta^{13}\text{C}$ response to fossil fuel combustion alone⁹ suggests that a linear approximation is reasonable over this period, and that the Table 2 regression coefficients (plus standard errors) are directly comparable.

Assuming the results obtained by Keeling *et al.*²¹ represent an ubiquitous $\delta^{13}\text{C}$ trend in the free atmosphere (supported by their comparison of North American and South Pole $\delta^{13}\text{C}$ measurements in 1978), then only the Rebello and Wagener tree can have responded in a direct manner.

The differences between $^{13}\text{C}/^{12}\text{C}$ trends in trees from different regions, coupled with the differences between most $^{13}\text{C}/^{12}\text{C}$ trends in tree-rings representing the past 20 yr and the direct atmospheric measurement, seriously question the interpretation of any tree record in terms of global atmospheric behaviour, and in particular the implicit assumption of constant fractionation in the atmosphere-biosphere exchange. An explanation involving large scale modification of atmospheric $^{13}\text{C}/^{12}\text{C}$ ratios as a result of plant respiration²³, apparently conflicts with well known rates of mixing of CO_2 in the atmosphere²⁴ during photosynthesizing hours. A solution must be sought in the mechanism of fractionation in plants.

I thank Dr Don Adamson, Dr John Ogden, Trevor Bird and George Dolezal for dated timber and detail on the individual trees, John Ogden and Drs Graeme Pearman and Paul Fraser for discussions, and Geoff Richards for the bulk of the chemical preparation.

Note added in proof: With respect to the suggestion²⁵ that isotope variations might primarily reflect ring width changes, there is a general absence of correlation between $\delta^{13}\text{C}$ and ring width in these trees, apart from exceptions mentioned in the text.

Received 28 August 1980; accepted 14 January 1981.

1. Freyer, H. D. & Wiesberg, L. *Naturwissenschaften* **60**, 517–518 (1973).
2. Farmer, J. G. & Baxter, M. S. *Nature* **247**, 273–275 (1974).
3. Rebello, A. & Wagner, K. *Proc. int. Symp. on Environmental Biogeochemistry* Vol. 1 (ed. Nriagu, J. O.) 13–23 (Ann Arbor, Michigan, 1976).
4. Galimov, E. M. *Proc. int. Symp. on Environmental Biogeochemistry* Vol. 1 (ed. Nriagu, J. O.) 3–11 (Ann Arbor, Michigan, 1976).
5. Stuiver, M. *Science* **199**, 253–258 (1978).
6. Wilson, A. *Nature* **273**, 40–41 (1978).
7. Farmer, J. G. *Nature* **279**, 229–231 (1979).
8. Freyer, H. D. *Tellus* **31**, 124–137 (1979).
9. Tans, P. P. & Mook, W. G. *Tellus* **32**, 268–283 (1980).
10. Grinstead, M. J., Wilson, A. T. & Ferguson, C. W. *Earth planet. Sci. Lett.* **42**, 251–253 (1979).
11. Freyer, H. D. *WMO Project on Research and Monitoring of Atmospheric CO_2* Rep. No. 1, Appendix E (WMO Geneva, 1979).
12. Freyer, H. D. *Tellus* **31**, 308–312 (1979).
13. Pearman, G. I. *Clean Air* 20–26 (1977).
14. Ogden, J. *Aust. J. Ecol.* **3**, 339–356 (1978).
15. Ogden, J. *Tree-Ring Bull.* **38**, 1–13 (1978).
16. LaMarche, V. C. Jr., Holmes, R. L., Dunwiddie, P. W. & Drew, L. G. *Tree-ring Chronologies of the Southern Hemisphere* Vol. 4 (Laboratory of Tree-Ring Research, Arizona, 1979).

17. Pearman, G. I., Francey, R. J. & Fraser, P. J. B. *Nature* **260**, 771–773 (1976).
18. Fraser, P. J. B., Francey, R. J. & Pearman, G. I. *DSIR Bull.* **220**, 67–73 (1978).
19. Francey, R. J. *et al.* (in preparation).
20. Keeling, C. D. *Geochim. cosmochim. Acta* **24**, 299–313 (1961).
21. Keeling, C. D., Mook, W. G. & Tans, P. P. *Nature* **277**, 121–123 (1979).
22. Keeling, C. D., Bacastow, R. B. & Tans, P. P. *Geophys. Res. Lett.* **7**, 505–508 (1980).
23. Mazany, T., Lerman, J. C. & Long, A. *Nature* **287**, 432–435 (1980).
24. Keeling, C. D. *Geochim. cosmochim. Acta* **13**, 322–334 (1958).
25. Wigley, T. M. L., Gray, B. M. & Kelly, P. M. *Nature* **271**, 92–93 (1978).

Ixtoc 1 oil spill: flaking of surface mousse in the Gulf of Mexico

John S. Patton & Mark W. Rigler

Department of Microbiology, University of Georgia, Athens, Georgia 30602, USA

Paul D. Boehm & David L. Fiest

Energy Resources Co. Inc., 185 Alewife Brook Parkway, Cambridge, Massachusetts 02138, USA

The blowout at the Ixtoc 1 offshore oil rig in the Gulf of Mexico (19°24' N, 92°12' W) produced a major oil spill. An estimated 140 million gallons of oil escaped before the well was finally capped¹; this compares with a spill of 66 million gallons of oil from the tanker *Amoco Cadiz*^{2,3}. In contrast to the rapid, close-to-shore release of the tanker's cargo, the Ixtoc 1 well released oil for 9 months into the open ocean where winds and currents dispersed the floating mousse (a light brown and orange water-in-oil emulsion) which had formed at the wellhead. Two and a half months after the blowout the western boundary current in the Gulf of Mexico^{4,5} had brought mousse to the American shore. An ultra-thin sheen also formed from the oil and was visible as a smoothness of sea-surface texture relative to the slightly rougher unpolluted waters. We report here that at a distance of 750–1,000 km from the well, Ixtoc 1 mousse occurred on sheen-covered sea surface primarily as gramme-sized pancakes and milligramme-sized flakes. Relative to pancakes, flakes were more dense, depleted in aliphatic compounds and enriched in polar compounds. Both particles were covered with a brittle 5–10 µm skin. Pancakes, skin, flakes and surface sheen seem to represent the four macroscopic component parts of a pelagic weathering cycle of Ixtoc 1 mousse.

Samples were collected from three stations near the South Texas coast and from one station near the wellhead (Table 1). Generally mousse particles occurred within or very close to areas covered with sheen, at an estimated concentration, based on scale-referenced sea-surface photographs, of 0.2–2.0 g mousse per m² of sea surface. Divers estimated that 99% of the mousse occurred on the sea surface. Flakes and pancakes did not always occur in the same sheen area. Flakes were collected with a stainless-steel screen and pancakes with a dip net. Station 1 was in a flake-pancake area of at least 20 km width, with flake concentrations ranging from 92 to 158 flakes m⁻² (470–806 mg flake m⁻²). This is ~400 times above the normal background tar concentration in that area of the Gulf⁶. Both types of mousse usually occurred in the shape of flattened disks with thickness usually less than one-fifth of the width. Samples were stored in filtered seawater in the dark at 4 °C in the presence of 0.02% sodium azide⁷. The visible production of thin sheen from mousse was not observed although at one point between stations, in the middle of a hot calm day, sheen was seen to be emanating from pancakes. Large numbers of flakes or pancakes in small volumes of water showed no tendency to aggregate. There was no size continuum between pancakes and flakes; the 1–4 mm thick flakes seemed to originate from the cracked scaly crust of the pancakes (Fig. 1a) and flakes could be produced by shaking pancakes in seawater.

Microscopic views of flakes are shown in Fig. 1b–h. Figure 1b shows the penetration of light through the edges of two flakes

observed with normal transmission light microscopy. Figure 1c, d shows the surface topography of two flakes viewed by reflected epifluorescence microscopy. Figure 1c shows the cracked surface of a flake that was touched by a Pasteur pipette. The light-coloured skin is shown on edge (S) and averaged 5–10 µm in thickness. The dark material within the cracks is emulsion (see also Fig. 1g). Figure 1d shows another flake with a more heterogeneous surface. The colour of the thin skin on an individual flake ranged from light yellow to brown. Often emulsion such as that in the upper-right corner of Fig. 1d was speckled with amorphous white fluorescent particles. Figure 1e shows a freeze-etched electron micrograph of the surface of flake skin with embedded diatom tests (D) and hundreds of flattened half-moon-shaped disks. Figure 1f is a blowup of the unidentified disks that are embedded flat and on end in the skin surface. Compared with flakes from Stations 2 and 3 flakes from Station 1 were highly contaminated with embedded material. Exposed emulsion on the surface of a flake is shown in Fig. 1g. Microbial colonies (M) were uncommon and usually occurred in isolated pockets on the flake. Loosely attached microbial cells and slime, such as on the edge of the flake in Fig. 1c, may have been more common but lost during sample transport. Microbial colonies were not observed on the brittle skin. Figure 1h shows a freeze-etched view of the interior of a flake that shows a smooth textured crude oil⁸ with a heterogeneous population of emulsified water droplets. The interior of flakes contained embedded particles such as dinoflagellate and diatom tests, fish scales and unidentified amorphous material. The pancake surfaces showed the same microscopic features as the flakes.

The pancake mousse exhibited gross chemical similarity to the wellhead mousse (Table 1), although their compositions were quite different: the latter containing significant quantities of saturated hydrocarbons in the *n*-C₁₀ to *n*-C₁₅ region (Fig. 2a) and large amounts of light aromatics (Fig. 2b), whereas the pancake mousse was almost totally depleted in alkanes eluting before *n*-C₁₅ and weathered such that only small levels of two ringed aromatics remained (Fig. 2c, d).

Mousse pancakes and their associated flakes were also compositionally different, (Table 1): the flakes were depleted in saturated hydrocarbons (*f*₁) and enriched in polar (*f*₃) content relative to the pancakes. Flakes were also more dense than their companion pancakes (Table 1) which could reflect a greater volume concentration of surface-embedded material than with the pancakes. Gas chromatograms of the mousse pancakes and flakes (Fig. 2c–e) indicate subtle but significant compositional differences between the two Ixtoc oil forms. The saturated (aliphatic) fraction of the flakes is more depleted in the lower boiling *n*-alkanes and reveals an unresolved complex mixture (UCM). Quantities of residual aromatics (Fig. 2d, f), consisting mainly of alkylated phenanthrenes and dibenzothiophenes, were reduced in the flakes although compositionally the two forms were similar. Figure 3 summarizes the hydrocarbon composition of wellhead mousse, Texas mousse pancakes and Texas mousse flakes. The presence of alkylated naphthalenes in the pancakes and their absence in the flakes indicates differences in the levels of the more soluble and potentially toxic naphthalenes in the mousse.

Changes in mousse composition can be attributed to rapid evaporation and dissolution⁹ in the early weathering stages and to photooxidation and evaporation in the later stages¹⁰. Differences between flake and skin compositions, reflected in severe depletion of aromatic hydrocarbons in the latter (compare Fig. 2f with h) are probably attributable to the importance of photooxidation in the skin. The high density of skin (>1.022 g cm⁻³) (Table 1) was due to a high proportion of embedded material.

The world oceans are widely polluted with floating petroleum residues—pelagic tar^{11–15}. Approximately one-third of the pelagic tar samples from the Gulf of Mexico¹⁶ and more than half of those from the central North Atlantic¹⁷ exhibit biomodal paraffin distributions in their gas chromatograms which indicate that the residues come from tanker sludge¹⁸. The remaining

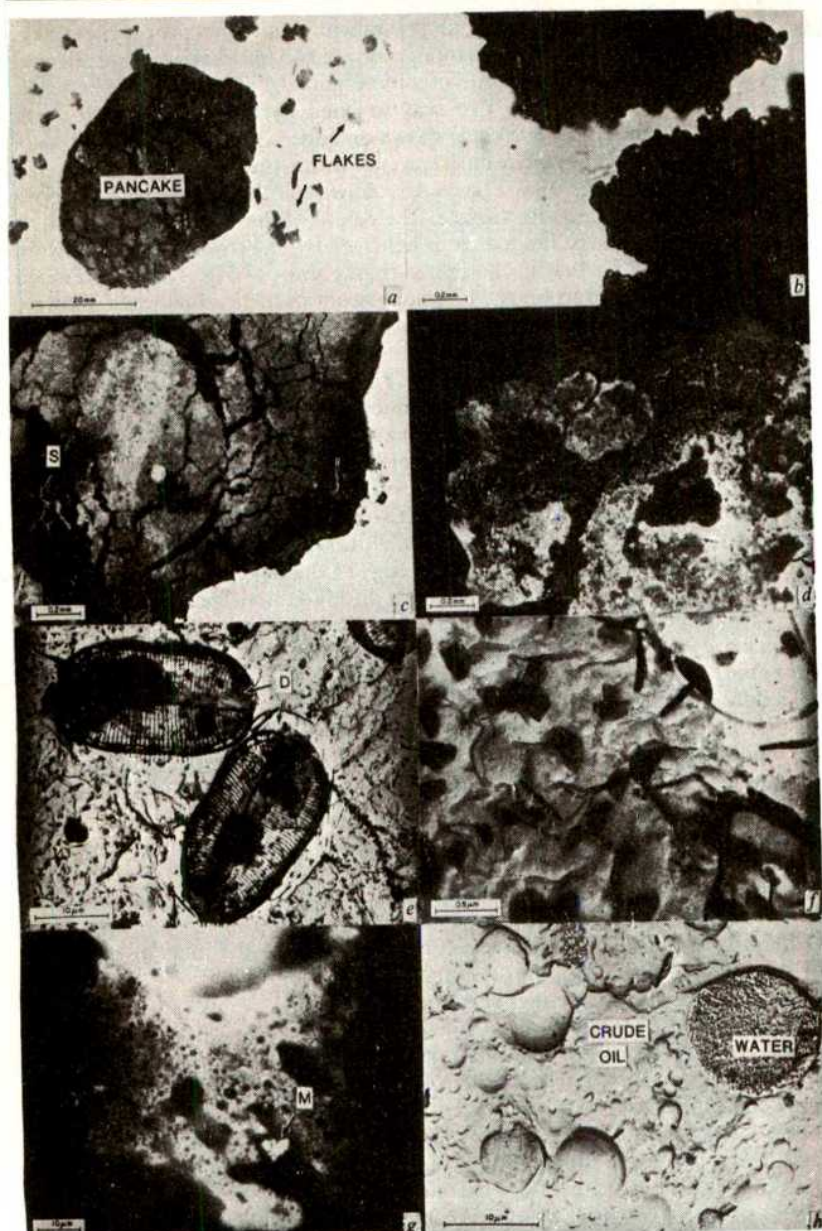


Fig. 1 A microscopic view of mousse flakes produced by weathering of Ixtoc 1 mousse pancakes in the Gulf of Mexico. *a*, A typical pancake and numerous flakes taken at Station 2 from the sea surface. *b*, Tungsten light penetration of two flakes from Station 2. *c*, Epifluorescence microscopy of flake surface showing skin thickness at S and semi-circular pipette imprint. *d*, Epifluorescence microscopy of flake from Station 1 showing dark regions not covered with light coloured skin but speckled with amorphous fluorescent material. *e*, Fortuitous freeze-etched electron micrograph of Station 1 flake skin with embedded diatom tests (D) and hundreds of tiny plates shown at higher magnification in *f*. *g*, Oil immersion epifluorescence microscopy of Station 1 flake surface showing exposed emulsion and a microbial colony (M) in a crevice of the flake. *h*, Freeze-etched electron micrograph of the interior of a Station 1 flake.

Table 1 Physical and chemical characteristics of Ixtoc 1 mousse taken from the surface of the Gulf of Mexico at different times and locations following the 3 June 1979 blowout

Station*	Sample	Distance from wellhead (km)	Maximum time of weathering (days)	Average weight (mg)	Average density (gm cm ⁻³)§	Chemical composition (% total)		
						Aliphatics <i>f</i> ₁	Aromatics <i>f</i> ₂	Polars <i>f</i> ₃
1	Flakes	1,025	79	5.1 (<i>n</i> = 122)	1.017	49.9	16.1	34.0
2	Flakes	875	75	3.5 (<i>n</i> = 44)	1.014	48.6	23.1	28.4
3	Flakes	758	77	1.7 (<i>n</i> = 171)	1.011	44.0	23.2	32.7
					Mean ± s.d.	47.5 ± 3.1	20.8 ± 4.1	31.7 ± 2.9
1	Pancakes	1,025	79	4200 (<i>n</i> = 15)	1.015	56.5	27.7	15.7
2	Pancakes	875	75	87,000 (<i>n</i> = 6)	1.011	58.7	16.5	24.8
					Mean ± s.d.	57.6 ± 1.6	22.1 ± 5.6	20.2 ± 6.4
1	Pancake skin†	1,025	79	<0.1	1.022	50.6	19.0	30.4
4	Wellhead mousse	7	7	—	—	55.7	25.0	19.3
4	Wellhead oil-in-water dispersion‡	<1	<1	—	—	49.3	42.6	8.1

* Station 1: 21 August 1979, 27°29' N, 96°44' W; Station 2: 17 August 1979, 26°10' N, 96°24' W; Station 3: 19 August 1979, 24°10' N, 97°37' W; Station 4: 17 September 1979, 19°24' N, 92°12' W.

† Pancake skin corresponds to the 5–10 µm skin shown in Fig. 1c, d. It was isolated from pancakes by shaking them in seawater. Flakes produced floated with the pancakes while the tiny pieces of skin sank to the bottom, indicating a density greater than seawater (>1.022).

‡ Oil collected at the wellhead with a GO-Flo bottle before it had reached the surface ('reference' wellhead mousse).

§ Density was measured by placing samples in solutions of different salinities. All samples sank at salinities 15‰ (1.009 g cm⁻³).

|| Fractions were isolated by silica gel column chromatography²⁷: Aliphatics-hexane eluate; aromatics-hexane/methylene chloride (50:50) eluate; polars-methanol eluate.

pelagic tar has a single paraffinic maximum in its gas chromatogram which is thought to be typical of weathered crude oils^{16,17}. The distinction between this type of pelagic tar and Ixtoc 1 mousse which also exhibits a single paraffinic maxima (Fig. 2c) is not clear. The presence of a stable water-in-oil emulsion seems to characterize mousse. Mousse pancakes and flakes from Stations 1 and 2 averaged 45–53% water as determined by centri-

fugation (after heating the mousse in a boiling water bath for 10 min). Water contents of pelagic tars have rarely been measured but presumably the waxy tanker precipitates contain little if any water, and the other so-called weathered crudes with single paraffin maxima may have been emulsion or mousse. It is not known whether mousse can become non-emulsified pelagic tar nor how crude oil composition effects emulsion formation

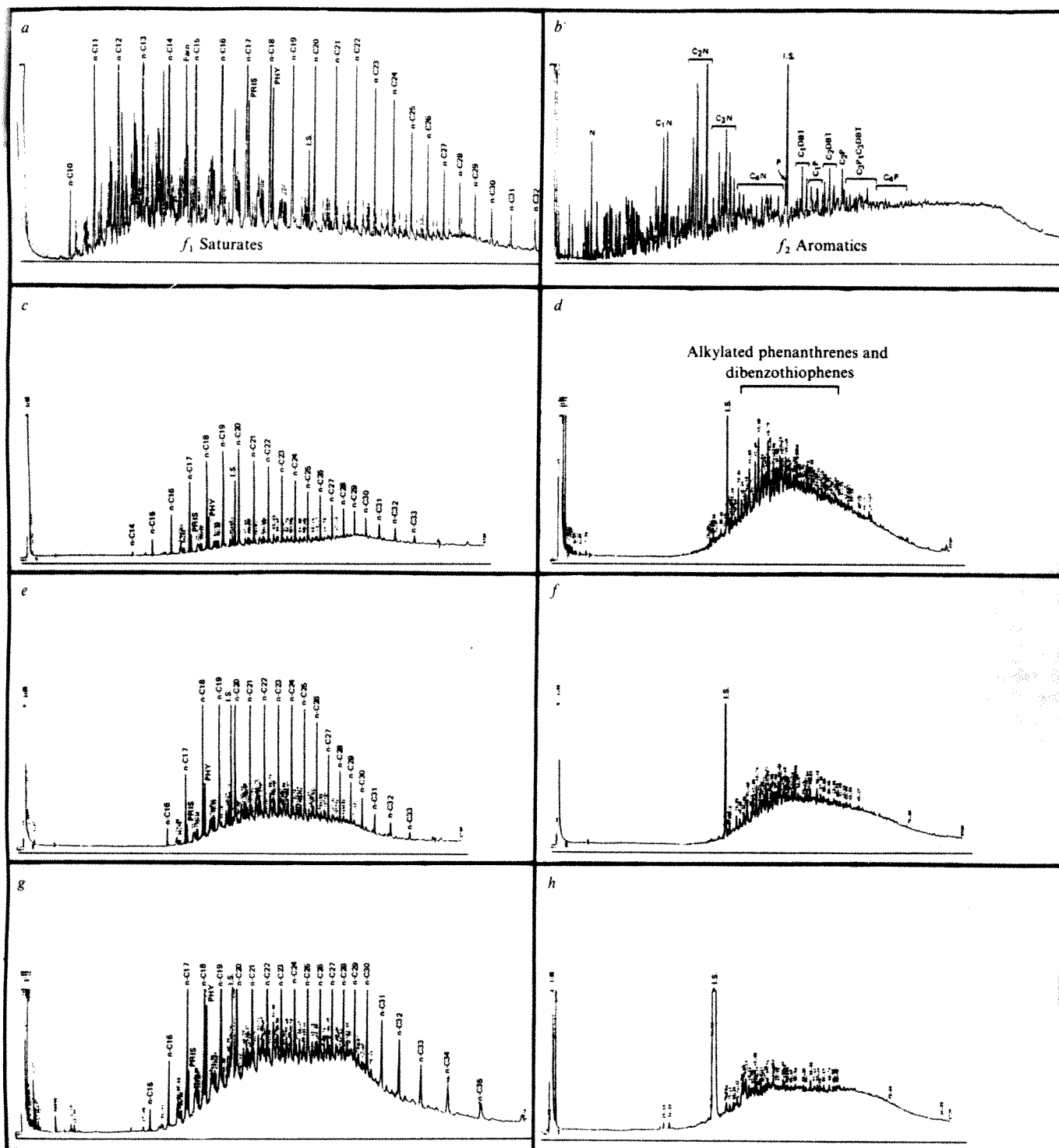


Fig. 2 Glass capillary gas chromatograms of f_1 , saturated (left column) and f_2 aromatic (right column) hydrocarbons from different Ixtoc 1 mousse fractions. An unresolved complex mixture (UCM) appears as a hump below the peaks in most of the chromatograms. *a, b*, Fractions from 'reference' wellhead oil obtained 0.4 km from the blowout (n -C₁₀, n -C₁₁, and so on represent n -alkanes; farn, farnesane; pris, pristane; phy, phytane; N, naphthalene; P, phenanthrene; DBT, dibenzothiophene C₁N, C₂N, C₁P, C₂P, C₁DBT and so on represent alkyl-substituted aromatics; GC, Hewlett Packard 5840A, 30-m SE-20 fused silica (J & W Scientific) splitless injection; temperature, 40–275 °C at 3° min⁻¹. Helium carries gas 1 cm³ min⁻¹ (*c, d*). Gas chromatograms of mousse pancakes obtained at Station 2 along the Texas coast ~875 km from the blowout site (see Table 1). IS, internal standard; (f_1 , androstane; f_2 , deuterated anthracene). *e, f*, Gas chromatograms of mousse flakes from Station 2. *g, h*, Gas chromatograms of Station 1 skin (see Table 1).

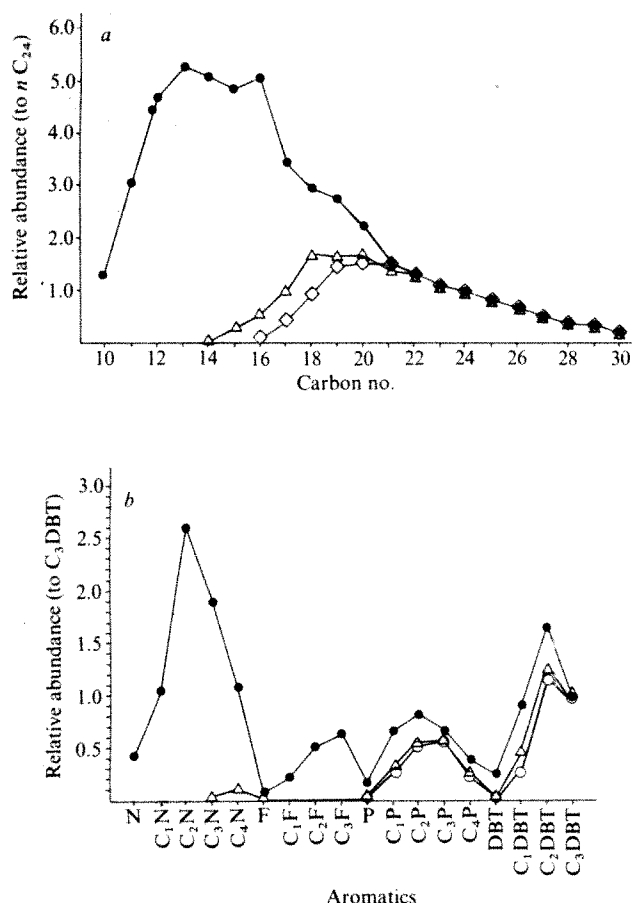


Fig. 3 Relative compositions of wellhead mousse, Texas mousse pancakes and mousse flakes: a, saturates; b, aromatics. ●, Wellhead mousse; △, mousse pancakes; ○, mousse flakes.

and stability; some crude oils may even be incapable of forming pelagic tars because of low asphaltene content¹⁶.

Our microscopic examinations suggest that biological factors are not directly involved in causing Ixtoc 1 mousse to crack and flake. Extensive microbial colonization could have been inhibited by the rate of flaking, by photochemical oxidation products^{19,20} or by the lack of nutrient²¹ in the surface waters of the Gulf. Evaporative weathering and physical fragmentation are thought to be the principal processes affecting pelagic tar residues at sea²²⁻²⁴ and wave motion has been postulated as a mechanism for mousse fragmentation^{25,26}. Our observations suggest that weathering processes first form a cracked, scaly surface on pancakes which then 'flake' in turbulent sea. The flaking process exposes fresh mousse within the pancakes and this relatively unweathered mousse creates surface sheen which emanates from the pancake until a skin forms. The skin flakes off on agitation thus ultimately breaking down the mousse pancake as well as altering its composition.

Flaking of stable, water-in-oil emulsion particles seems to be a significant intermediate in the open-sea dispersion of the Ixtoc 1 oil spill. The process causes an increase in the surface-to-volume ratio of spilled oil which would accelerate surface-dependent evaporative dissolution and photochemical weathering and enhance the rate of microbial colonization. The increased density of the flakes relative to pancakes means that their lifetime in the water column may be shortened. The persistence of floating pancakes may be short-lived because of flaking.

We thank John Robinson, John Farrington, Donald Atwood, Pat Parker, Tony Amos and the scientists and crews of the RV Longhorn and RV Researcher for encouragement, constructive criticism and support. This work was supported by NOAA and a small grant from the Georgia Research Foundation.

Received 20 August; accepted 16 December 1980.

1. *Oil Spill Intelligence Reports* Vol. 2, No. 22 (1979)–Vol. 3 No. 13 (1980).
2. Spunner, M. F. *Mar. Pollut. Bull.* **9**, 281–284 (1978).
3. Butler, J. N. *Ocean. Ind.*, October 101–112 (1978).
4. Sturges, W. & Blaha, J. P. *Science* **192**, 367–369 (1976).
5. Molinari, R. L. *J. phys. Oceanogr.* **8**, 779–784 (1978).
6. Pequegnat, L. H. *Contr. mar. Sci.* **22**, 31–39 (1979).
7. Kleinhofs, A., Owais, W. M. & Nilan, R. A. *Mutat. Res.* **55**, 165–195 (1978).
8. Eley, D. D., Key, M. H., Symonds, J. D. & Willison, J. H. M. *J. Coll. Interface Sci.* **54**, 462–467 (1976).
9. Boehm, P. D., Fiest, D. L., Mackay, D. & Paterson, S. in *Proc. 1981 Oil Spill Conf.* (American Petroleum Institute, 1981).
10. Overton, E. B. et al., in *Proc. Conf. on the Preliminary Scientific Findings of the Researcher/Pierce cruise to the Ixtoc 1 Blowout* (NOAA, Washington DC).
11. Horn, M. H., Teal, J. M. & Backus, R. H. *Science* **168**, 245–246 (1970).
12. Heyerdahl, T. *Biol. Conserv.* **3**, 164–167 (1971).
13. Zolsnay, A. *Mar. Pollut. Bull.* **8**, 116–117 (1970).
14. *Petroleum in the Marine Environment*, 48 (National Academy of Sciences, Washington DC, 1975).
15. Eagle, G. A., Green, A. & Williams, J. *Mar. Pollut. Bull.* **10**, 321–325 (1979).
16. Jeffrey, L. M., Peguegnat, W. E., Kennedy, E. A., Vos, A. & James, B. M. *NBS Spec. Publ. Marine Pollution Monitoring* **409**, 233–235 (1974).
17. Butler, J. N. & Harris, J. C. *Mar. Chem.* **3**, 1–7 (1975).
18. Butler, J. N., Morris, B. F. & Sass, J. *Spec. Publ. Bermuda Biological Station for Research* **10**, 1–346 (1973).
19. Larson, R. A., Bott, T. L., Hunt, L. L. & Rogenmuser, K. *Envir. Sci. Technol.* **13**, 965–969 (1979).
20. Hansen, H. P. *Mar. Chem.* **3**, 183–188 (1975).
21. Perry, J. J. *Microbiol. Rev.* **43**, 59–72 (1979).
22. Wade, T. L. & Quinn, J. G. *Mar. Pollut. Bull.* **6**, 54–57 (1975).
23. Butler, J. N. *Mar. Chem.* **3**, 9–21 (1975).
24. Morris, B. F., Butler, J. N., Sleeter, T. D. & Cadwallader, J. *Rapp P.-V. Reun. Cons. Int. Explor. Mer.* **171**, 107–116 (1977).
25. Audunson, T. *Mar. Envir. Res.* **3**, 35–61 (1980).
26. Forrester, W. D. J. *Mar. Res.* **39**, 151–170 (1971).
27. Farrington, J. W. & Tripp, B. W. in *Marine chemistry in the Coastal Environment* (ed. Church, T. M.) 267–284 (American Chemical Society, Washington DC, 1975).

Annual laminations in the sediments of Loe Pool, Cornwall

H. L. K. Simola

Karelian Institute, University of Joensuu, SF-80101 Joensuu 10, Finland

M. A. Coard & P. E. O'Sullivan

Department of Environmental Sciences, Plymouth Polytechnic, Drake Circus, Plymouth, Devon PL4 8AA, UK

Annually laminated freshwater lacustrine sediments have been recorded at several sites in central Europe¹⁻⁵, North America⁶⁻¹³ and Fennoscandia¹⁴⁻²⁴. The presence of laminations may reflect either (1) regular changes within the lake ecosystem itself or (2) variation in the intensity of erosion and transport of material from the catchment, particularly where instability in the lake-watershed system has occurred as a result of human activities²². The principal cause of lamination is, therefore, seasonal variation of environmental conditions, particularly climate. Lakes with laminated sediments tend to be physically deep, exhibit a strong seasonal stratification, and be situated in areas of continental climate. We describe here what we believe to be the first reported instance of a long sequence of laminated lake sediments from Great Britain. Unlike most of the previous examples, these have been formed in a shallow, polymictic lake, in an oceanic climate.

Loe Pool (Fig. 1) is a eutrophic freshwater lagoon at ~4 m OD, 1 km south of Helston in south-west England (GR SW 648250, lat 50°4' N, long 5°17' W) with an area of ~44 hectares and a mean depth of 4 m. Its catchment covers ~50 km², and is mainly farmland, with one major settlement (Helston, population ~10,000). The main stream entering the pool is the River Cober, which drains ~90% of the total catchment. In the nineteenth and early twentieth century this area was the site of extensive mining operations, principally for tin²⁵.

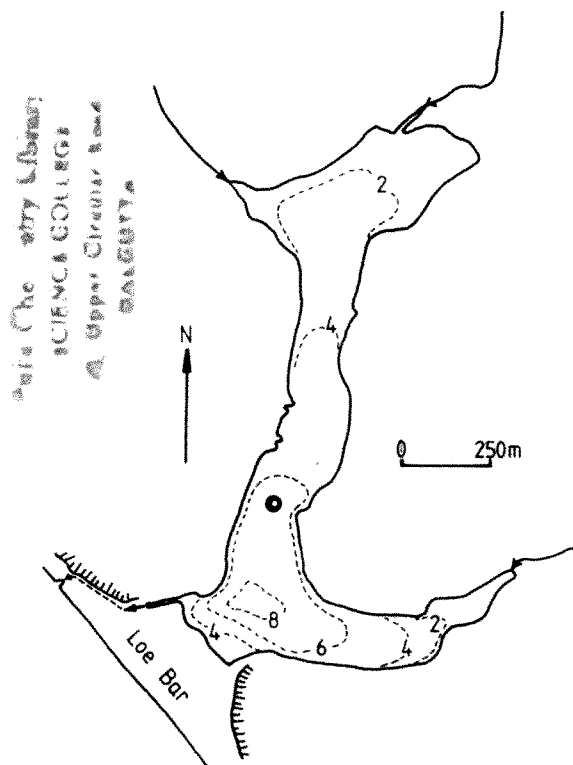


Fig. 1 Loe Pool, Helston, Cornwall, showing depth contours (dashed line) and position of coring site (●).

Three sediment cores were taken from beneath ~7 m of water at the location shown in Fig. 1, where studies (M.A.C. unpublished data) indicate that an undisturbed conformable sequence of sediments is available. One was obtained using a 3-m version of the Mackereth corer²⁶ and two with 1-m *in situ* freezing device²⁷. The Mackereth core was frozen at ~-20 °C to preserve lamination. The surfaces of all cores were then cleaned while still frozen to facilitate description and photography of sediment structures. The stratigraphy of the three cores was easily matched by using prominent marker horizons and characteristic sequences of laminations.

The cores were then allowed to dry in the freezer for a few days to allow adhesive tape preparations to be made¹⁸. These enable the fine structure of the sediment to be examined,

especially for changes in abundance of sub-fossil diatoms. The stratigraphy of the sediment at the sampling site was:

0–40 cm: Highly organic *gyttja*, with four or five pairs of light and dark brown laminations just below the sediment surface (0–5 cm), the rest more homogeneous.

40–120 cm: Irregularly laminated sequence containing red and grey clays alternating with darker layers.

120–300 cm: More regularly laminated sediment consisting of paired grey and black layers, average thickness 3 cm per pair.

The sections 0–7.2 cm, 40–110 cm, 194–206 cm and 230–242 cm were examined. Remains of diatoms, other algae and vivianite crystals were recorded in consecutive 200-μm fields. We shall concentrate here on information obtained from the first and third sections, results being presented in Figs 2 and 3.

In all sections studied, diatom and other algal taxa appear in repeating sequences, the cyclic nature of which we attribute to seasonal algal production and sedimentation. It is therefore possible to define annual increments of deposition, which correspond closely with clearly visible stratigraphic changes. We can therefore identify laminations in these sediments which are truly annual in their nature. Interruptions to the sequence are mainly associated with the deposition of layers of clay, sometimes massive, in which diatoms are relatively scarce, and which we consider to represent dilution of diatom influx by allochthonous, clastic material.

For example, in Fig. 2 (0–7.2 cm), six phases (a–f) are defined. In each of these, *Thalassiosira pseudonana* and *Cyclotella meneghiniana* are succeeded by *Asterionella formosa*, *Melosira granulata* var. *angustissima* and then *Melosira varians*. In monthly plankton sampling of the pool by one of us (M.A.C.) in 1979, the first two taxa were prominent during spring and early summer, and the *Melosira* species in the autumn and winter. It is therefore considered that this sedimentary sequence records seasonal diatom succession in the pool, the *Thalassiosira*/*Cyclotella* stages representing the spring and summer, the *Asterionella*/*Melosira* stages the autumn and winter. The appearance of *Fragilaria* spp. and 'other' (mainly sessile) diatoms in layers thought to represent the winter, would then be consistent with the idea of increased influx of sediment from the littoral and the catchment during these months. Similarly, the summer abundance of *Pediastrum boryanum* in lake plankton samples is recorded in summer laminations. Crystals of vivianite ($\text{Fe}_3(\text{PO}_4)_2 \cdot 8\text{H}_2\text{O}$) seem to be concentrated in some winter layers. This has also been observed in Lovöjärvi (Finland)^{17,18}.

In this section, summer layers coincide with the lighter bands of sediment, and winter layers with the darker. Below 5 cm, however, the sediment structure and the diatom peaks are less

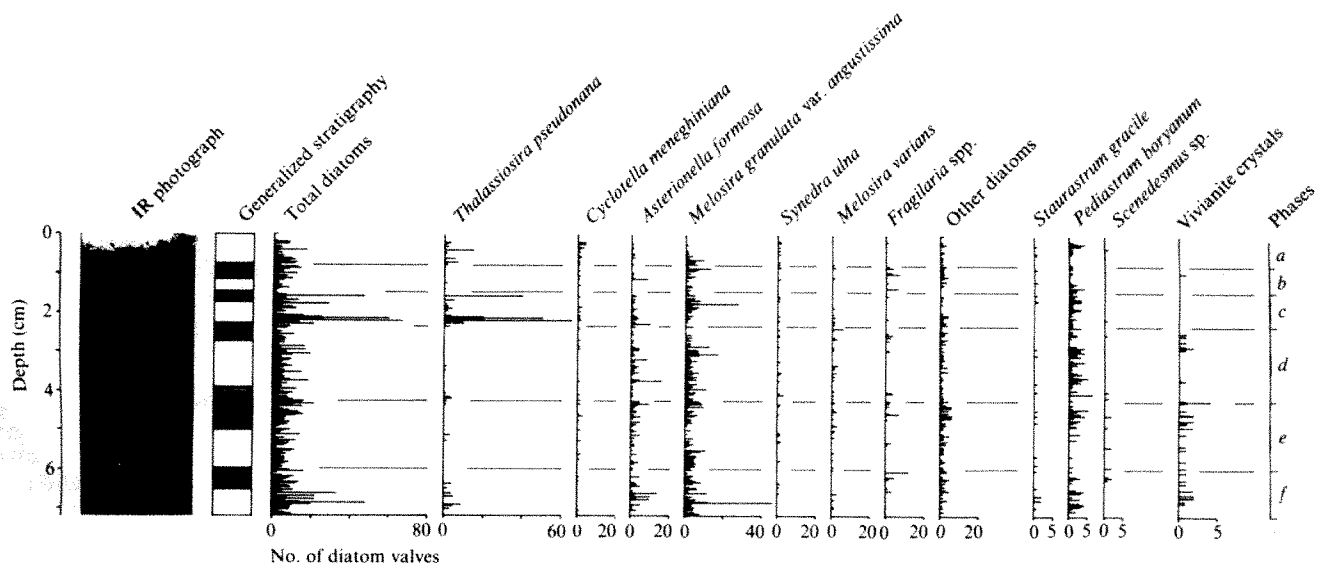


Fig. 2 IR photograph of and results of diatom, other algal and vivianite crystal counts from the section 0–7.2 cm of frozen sediment from Loe Pool.

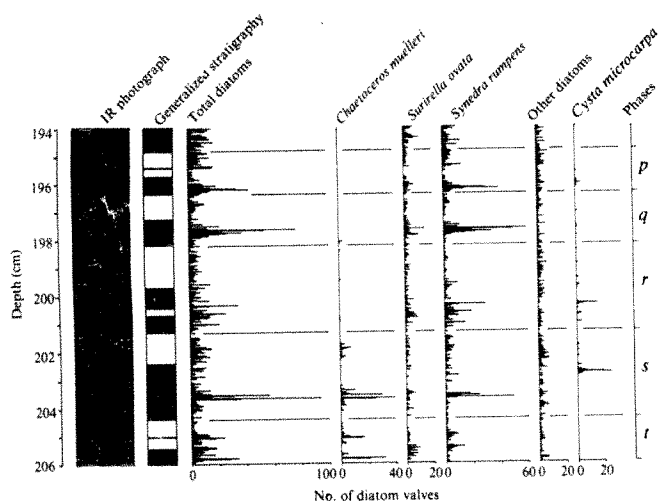


Fig. 3 IR photography, and results of diatom and *Cysta microcarpa* counts from the section 194–206 cm of a frozen sediment core from Loe Pool. Irregular, diagonal striations are caused by ice-crystal formation during freezing of the Mackereth core.

well defined. Inspection of these and other cores indicates considerable bioturbation of the sediments between 5 and 40 cm.

Figure 3 (194–206 cm) shows a section in which five peaks of diatom deposition (*p–t*) are separated by relatively barren layers. The most abundant taxa present are *Chaetoceros muelleri*, *Surirella ovata* and *Synedra rumpens*, which tend to appear in a sequence which denotes seasonal succession. On three occasions, maxima of these planktonic diatoms are immediately followed by peaks of *Cysta microcarpa* (*sensu* Nygaard)²⁸. We interpret the planktonic diatom peaks, which coincide with darker bands of sediment, as representing the summer months, and the relatively barren layers, which correspond to the grey-brown clays, and in which 'other' diatoms largely occur, as the winter.

Figure 4 clearly shows that despite the oceanicity of the climate of this part of Europe, a pronounced winter maximum of rainfall and stream flow occurs, and that summer temperatures substantially exceed those of the winter. This contrast between the seasons seems sufficient to account for the formation of the Loe Pool laminations.

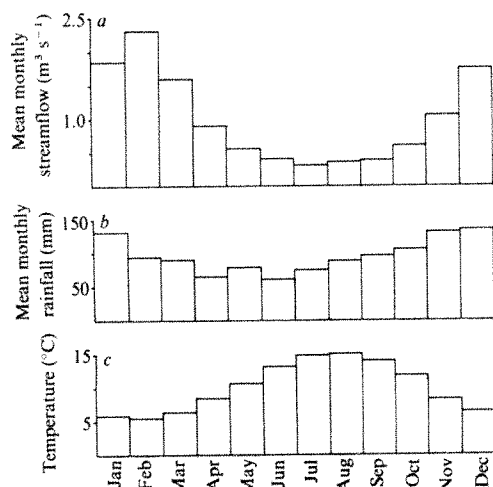


Fig. 4 *a*, Mean monthly streamflow (River Cober, 1970–79); *b*, mean monthly rainfall (1941–70); and *c*, mean monthly temperature (1960–74) for the Loe Pool catchment. Streamflow data from the South West Water Authority gauging station at Helston, rainfall data from Wendron, and temperature data from RNAS Culdrose.

From 300 to ~120 cm these consist of pairs of black (summer) and grey-brown (winter) layers, recording increased inflow of clastic material in the winter months, covering a period of ~60 yr. Then, between 120 and 40 cm, massive clay layers occur, which, according to diatom analysis not described in detail here (H.L.K.S. and M.A.C., in preparation), cover a period of only 7 yr. Here a high proportion of the sediment consists of material originating as mining wastes, and peaks in planktonic diatoms are separated by as much as 20 cm of clay in which diatom influx has been considerably diluted, thus indicating very rapid accumulation. We correlate this sequence with the most recent period of active mining in the catchment which ended in AD 1938. Within this section, two layers of red, haematite-rich clay occur, the upper at 40–46 cm, the lower at 65–70 cm. Documentation²⁹ shows that during the 1920s and 1930s the pool was heavily polluted by mining wastes, often to the extent that discoloration of the waters occurred. We therefore conclude that the red clay layers were deposited during this time. Between 1860 and 1920 active, but less intensive mining

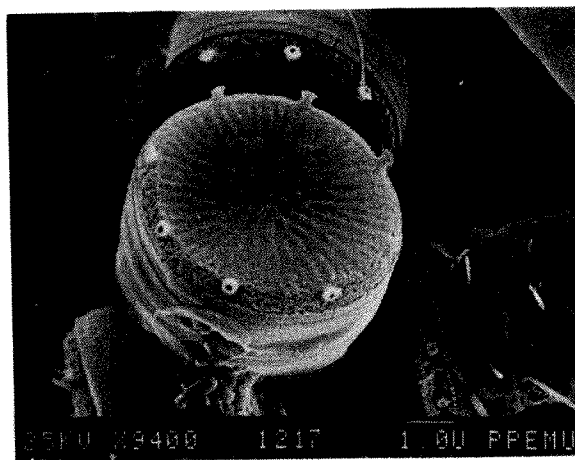


Fig. 5 Scanning electron micrograph of *Thalassiosira pseudonana* Hasle et Heimdal from Loe Pool plankton samples. (Photograph Martin Coard.)

took place²⁵. Above 40 cm, laminations are not observed until ~5 cm from the sediment surface. Here there are four or five laminations which are not visually prominent, but which are well defined in the algal stratigraphy. Between 5 and 40 cm the sediment is considerably bioturbed, but ¹³⁷Cs analysis shows that the 1963 peak lies between 14 and 16 cm depth. This suggests a mean accumulation rate for the section between 16 cm and the sediment surface of 1 cm yr⁻¹, which is in agreement with the rate calculated from lamination counts just below the sediment surface.

The section 0–40 cm, between the top of the upper haematite clay and the sediment surface, thus represents the period since 1938, and has a mean accumulation rate very close to 1 cm yr⁻¹. Together with a sedimentation rate of 80 cm in 7 yr for the section 40–120 cm (based on diatom stratigraphy), and of 3 cm yr⁻¹ for the section 120–300 cm (based on lamination counts), this estimate allows us to conclude that the Loe Pool cores cover a period of ~110 yr.

Between ~1870 and ~1920 the main cause of lamination seems to have been a steady inflow of mining wastes in winter. During the period of most intensive mining (1920–38) this increased to massive proportions (80 cm in 7 yr). Between 1940 and 1975 laminations were not formed. The presence of the remains of numerous burrows indicates that this was a phase of considerable benthic activity. The formation of laminations in very recent years is attributed to increased eutrophy of the pool, which has led to occasional instances of anoxia at the sediment

surface during the summer months. These most recent laminations are therefore being produced in different conditions from those formed at an earlier date, which may in part account for colour difference between the respective summer and winter layers.

We thank Lieutenant Commander J. P. Rodgers and the National Trust for permission to carry out work at Loe Pool. We also thank R. S. Cambray for ^{137}Cs analysis and Dr A. C. Hamilton for the loan of a Mackereth corer. The work was carried out while H.L.K.S. and M.A.C. were supported by Devon Local Authority Research Assistantships, and H.L.K.S. by a grant from the Alfred Kordelin Foundation (Finland).

Received 10 October 1980; accepted 14 January 1981.

1. Nipkow, F. Z. *Hydrol. (Budapest)* **1**, 1–23 (1920); *Schweiz. Z. Hydrol.* **4**, 71–80 (1927).
2. Thomas, E. A. *Vjschr. naturf. Ges. Zürich* **109**, 103–142, (1964).
3. Geyh, M., Merkt, J. & Müller, H. *Arch. Hydrobiol.* **69**, 366–399 (1971).
4. Thompson, R. & Kelts, K. *Sedimentology* **21**, 577–596 (1974).
5. Kelts, K. & Hsü, K. M. in *Lakes: Geology, Chemistry, Physics* (ed. Lerman, A.) (Springer, Berlin, 1979).
6. Tippet, R. *Can. J. Bot.* **42**, 1693–1709 (1964).
7. Ludlam, S. D. *Limnol. Oceanogr.* **14**, 848–857 (1969); **21**, 743–746 (1976).
8. Craig, A. J. *Ecology* **53**, 46–57 (1972).
9. Anthony, R. S. *Limnol. Oceanogr.* **22**, 45–54 (1977).
10. Swain, A. *Quat. Res.* **3**, 383–396 (1973); **10**, 55–68 (1978).
11. Edmondson, W. T. *Verh. int. Ver. Limnol.* **19**, 770–775 (1975).
12. Allison, D. E. *Limnol. Oceanogr.* **15**, 138–144 (1970).
13. Cwynar, L. C. *Can. J. Bot.* **56**, 10–21 (1978).
14. Digerfeldt, G., Battarbee, R. W. & Bengtsson, L. *Geol. Förs. Stockh. Förh.* **97**, 29–40 (1975).
15. Renberg, I. *Geol. Förs. Stockh. Förh.* **98**, 335–360 (1976); *Early Norrland* **11**, 63–92 (1972).
16. Kukkonen, E. & Tynni, R. *Acta bot. fenn.* **90**, 1–30 (1971).
17. Saarnisto, M., Huttunen, P. & Tolonen, K. *Ann. Bot. Fenn.* **14**, 35–45 (1977).
18. Simola, H. *Ann. Bot. Fenn.* **14**, 143–148 (1977).
19. Vuorinen, J. *Pol. Arch. Hydrobiol.* **25**, 443–451 (1978).
20. Tolonen, M. *Ann. Bot. Fenn.* **15**, 177–208, 209–222, 223–240 (1978).
21. Koivisto, E. & Saarnisto, M. *Geogr. Ann.* **A60**, 55–61 (1978).
22. Simola, H. *Holarctic Ecol.* **2**, 160–168 (1979).
23. Tolonen, K. *Boreas* **9**, 11–19 (1980).
24. Battarbee, R. W., Cronberg, G. & Lowry, S. *Hydrobiologia* **71**, 225–232 (1980).
25. Hamilton-Jenkins, A. K. *Wendron Tin* (Wendron Forge Ltd, Helston, 1978).
26. Mackereth, F. J. H. *Limnol. Oceanogr.* **3**, 181–191 (1958).
27. Huttunen, P. & Meriläinen, J. *Ann. Bot. Fenn.* **15**, 128–130 (1978).
28. Berg, K. & Petersen, I. *Folia limnol. scand.* **8**, 1–273 (1956).
29. *Western Independent* Penzance (2 February 1928).

Supply of iron and manganese to an anoxic lake basin

W. Davison

Freshwater Biological Association, Windermere Laboratory, Ambleside, Cumbria LA22 0LP, UK

In seasonally anoxic lakes, the dynamic cycling of iron and manganese makes possible the study of geochemical cycling and processes of mineral formation, while the rapid lake processes may be a model for the analogous but slower oceanic processes. We have now measured concentration–depth profiles for iron and manganese to assess the flux of these elements from the sediments. The seasonal accumulation of iron in lake water can be accounted for by supply from the sediment, but the manganese profiles suggest that the major source of manganese is dissolution from particular material sedimenting through the water column.

This research is part of a concerted programme^{1,2} on the geochemical cycling of iron and manganese in lakes. The characteristic features of the process in lakes subject to deep-water anoxia are well known^{3–5}. Although the observed distribution of the various forms of iron in such conditions has recently been related to individual rate processes on the basis of detailed measurements¹, no such treatment exists for manganese. However, the concentration–depth profiles usually have a significantly different shape from those of iron, and this has been attributed³ to manganese being more easily reduced. It has been assumed^{3,6,7} that manganese is supplied from the sediment in anoxic conditions. Golterman⁸ speculated that

manganese oxides may be reduced in the water column to Mn(II) with concomitant oxidation of Fe(II), and Delfino and Lee⁹ realized that dissolution or desorption processes in the water column could be significant to the generation of Mn(II). Despite detailed studies of manganese in seasonally anoxic lakes^{9–11} the roles of dissolution of particulate matter or release from the sediment have not been clarified. However, most workers have looked at ‘macro-profiles’—concentrations measured at depth intervals of 1 m or greater—whereas to study sediment–water interactions ‘micro-profiles’—concentrations measured at centimetre distances from the sediment—need to be known.

Water samples were collected from the seasonally anoxic Esthwaite Water^{1,4,5} at approximately monthly intervals during the summer of 1980. A pump incorporating a sediment detector¹² was used to collect samples for both micro- and macro-profiles. The samples were analysed for Fe(II) and Mn(II) by polarography¹³, and for total iron and manganese by acid digestion procedures^{14,15}.

Figure 1 shows the results with respect to distance from the sediment and time of the year. Regression analysis of the micro-profiles for both total and soluble manganese measured between 5 and 50 cm from the sediment showed that the change in concentration with respect to distance from the sediment (dc/dz) did not significantly differ from zero ($P>0.8$). The mean of the 11 values for dc/dz was $-0.87 \mu\text{g l}^{-1} \text{cm}^{-1}$ (range -1.96 to $+0.52$). The slopes calculated from the macro-profiles, 0.5–1.5 m from the bottom showed that the total and polarographic measurements gave similar results and the value of dc/dz was $\sim -0.5 \mu\text{g l}^{-1} \text{cm}^{-1}$.

In contrast to manganese there is no polarographically measurable Fe(II) during May and June and it is only just measurable in early July (Fig. 1). During this time the total iron in the water column increases substantially. This has been explained^{1,4,5} by assuming that the oxidized sediment layer, which during the winter months overlies the very reducing sediment, is itself reduced when the lake stratifies and so Fe(II) diffuses into the water column. The overlying water is not then entirely reducing and so Fe(II) is oxidized almost immediately, and the resultant Fe(III) therefore builds up in the water column. When Fe(II) does appear in the water column in July and August a greater concentration gradient exists in the 25 cm of water nearest the sediment. Using regression analysis to calculate the slope reveals that the flux calculated from the micro-profiles was 10–30 times greater than that estimated from the slope of the macro-profile at 0.5–1.5 m.

Davison *et al.*¹ have shown that the lateral flux of iron supplied from sediment at the same depth as the water must be taken into account. Moreover, towards the latter part of the season, dissolution of particulate iron falling through the water column also becomes important. Thus the macro-profile cannot be used for estimating the direct flux from the sediment. However, from Fig. 1 it seems that in the confined zone of the bottom 25 cm of water the chemical gradient from the sediment dominates the lateral component of water movement and so a gradient which reflects the sediment–water interaction is established.

The seasonal increase in the iron and manganese in the bottom waters of the lake varies little from year to year^{1,4,5}. For 1979, 1.73 and 1.34 tonnes of iron and manganese respectively had accumulated in the larger lake basin by the end of the season. Assuming that both elements may dissolve for a period of 12 weeks from all that sediment which may possibly become anoxic, that is below 8.5 m depth, the mean micro-profile value for 23 July 1980 and 14 August 1980 of $-82 \mu\text{g l}^{-1} \text{cm}^{-1}$ for iron, results in 2.20 tonnes supplied from the sediment. For manganese, however, a slope of $-1 \mu\text{g l}^{-1} \text{cm}^{-2}$ results in 0.027 tonnes, only 2% of the manganese found to have accumulated in the lake. Therefore although iron may be sediment derived, even if a previously measured manganese concentration gradient, which is 5 times the value used here, is considered¹², manganese is not. This is verified by the shape of the manganese profiles. The virtual invariance of concentration with respect to depth in the stratified bottom waters indicates a lack of any

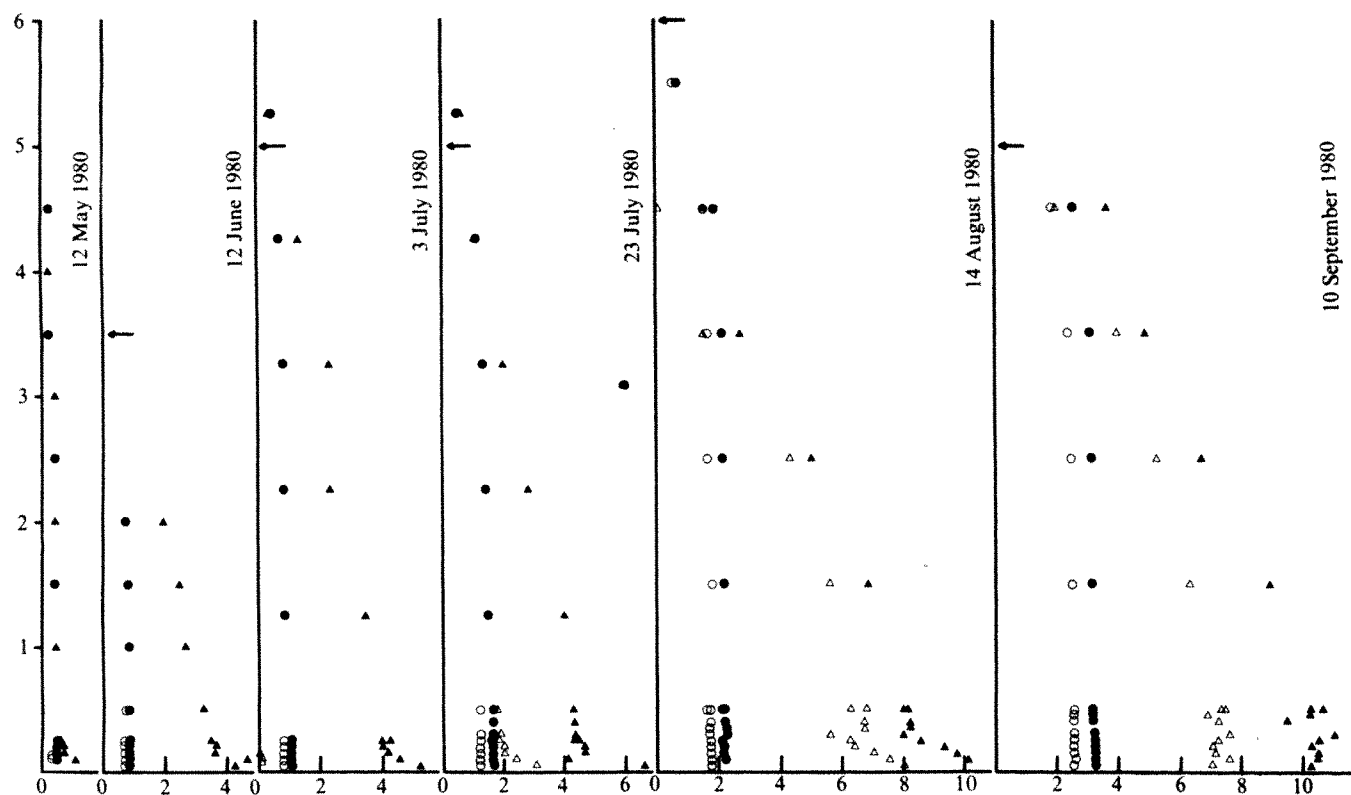


Fig. 1 The concentration of polarographically measured iron (Δ), total digested iron ($\blacktriangle$), polarographically measured manganese ($\circ$) and total manganese ($\bullet$) versus distance from the sediment. The position of the oxycline below which oxygen is not measurable ($<0.3 \text{ mg l}^{-1}$) by an *in situ* probe¹⁶ is indicated. On 12 May 1980 oxygen was still present at 4.8 mg l^{-1} near the sediment.

active point source or sink. Yet the progressive increase in concentration with time indicates a uniform supply of manganese at all depths. Comparison with the iron profiles shows that the shape cannot be accounted for by supply from lateral sediments; no similar increase in the waters in close proximity to the sediment is evident. On the macro-scale, supply from lateral sediment could not produce such a uniform profile because the ratio of sediment area-to-volume at any given stratum increases with depth.

Profile shapes show that at most 50% of the flux of iron at any depth is due to dissolution of particulate material¹. The calculation¹⁷ uses the concentration gradient and a knowledge of the vertical eddy diffusion coefficient to estimate the flux between each horizontal stratum of water which was assumed to be uniformly mixed. The lateral sediment derived flux was taken into consideration and any flux imbalance was attributed to *in situ* dissolution processes. Using the profiles of 14 August 1980 and 10 September 1980 for manganese the calculation suggests that dissolution accounts for 80–90% of the manganese flux. Although the approximations involved mean the results are semi-quantitative, the trend offers a possible explanation for the shape of the manganese profile. Particulate manganese from the inflowing streams must always be falling through the water column and may act as a source of supply for soluble manganese at all depths. The estimated 1.3 tonnes of manganese which accumulates in the hypolimnion results in a large increase in concentration in these bottom waters, but its contribution to the total manganese in the lake basin is not so great because of the relatively large volume of water in the surface layers. Before stratification uniform manganese concentration profiles with $100 \mu\text{g l}^{-1}$ of Mn could be measured, suggesting that 0.4 tonnes of manganese were present in the lake before any accumulation began. Moreover, it was estimated by gauging that $6 \times 10^6 \text{ m}^3$ of water flowed into the lake during June, July and August of 1980. The scant measurements indicate that the concentrations in the inflows are two to three times higher than in the epilimnion and so this volume of water probably carried ~ 1.2 tonnes of manganese into the lake. Therefore sufficient manganese must

either have been present in the lake or supplied to it, to account for the seasonal hypolimnetic increase.

The rapid increase in both polarographic and total manganese below the oxycline (Fig. 1) shows that the rate of vertical mixing in the metalimnion is not sufficiently rapid to mask the oxidation of the reduced form. Although the profiles of particulate manganese can be deduced from the difference between the two components plotted in Fig. 1, no conclusions can be drawn because the fraction of the particulate concentration which dissolves at any depth will depend on the rate of supply of manganese to the lake, the rate of dissolution, the sinking rate and the gradient of redox intensity. Simple solubility calculations show that neither MnCO_3 nor MnS can control the profile shape^{18,19}.

The build up of manganese in the hypolimnion therefore cannot be simply attributable to dissolution from the sediment. The shapes of the iron and manganese profiles show that simple vertical transport processes cannot account for the near-vertical manganese profiles and *in situ* supply by dissolution of sedimenting particulate material accounts for the observations.

The present results agree well with those of related studies. Eaton²⁰ showed that a considerable flux of the manganese in a fjord resulted from the dissolution of manganese oxides in anoxic conditions. There is evidence from estuarine work^{21–23} for manganese redissolving in the upper estuary where low pH and Eh prevail. The manganese profile in the permanently anoxic Black Sea has been modelled on the supposition of *in situ* dissolution²⁴. The same explanation may apply to a comparable peaked profile reported for a lake²⁵; slight peaks have also been reported for other lakes^{3,26}.

The interpretation of the iron and manganese ratios in the sediments of Esthwaite Water¹¹ has been based on the supposition that a substantial manganese fraction is lost from the sediment during anoxia: these results may require fresh interpretation. Care must be taken in extrapolating the results from artificially enclosed sediment–water systems to lakes^{4,5,27,28} because the possible input of manganese from dissolution of sedimenting particles is excluded. Mortimer^{4,5} reported the

presence of soluble manganese 40 days after Fe(II) had appeared in his laboratory tank experiments, which contrasts with the natural system where soluble manganese appears first.

Interstitial water profiles in sediments have been commonly used to estimate fluxes to the water column²⁸. The rate of transport is usually $\sim 10^3$ times slower than in the hypolimnion studied here. As micro-profiles discriminating to <25 cm are necessary for the hypolimnetic waters, by analogy the resolution of interstitial water measurements needs to be to <0.25 mm because lateral processes such as interchange with labile mineral fractions can still occur. Therefore caution should be exercised in using interstitial water profiles to estimate fluxes through the sediment-water interface.

I thank E. Tipping for constructive criticism and C. Woof for technical assistance.

Received 13 October 1980; accepted 14 January 1981.

1. Davison, W., Heney, S. I., Talling, J. F. Rigg, E. *Schweiz Z. Hydrol.* (in the press).
2. Tipping, E. *Geochim. cosmochim. Acta* **45**, 191–199 (1981).
3. Hutchinson, G. E. *A Treatise on Limnology* Vol. 1 (Wiley, New York, 1957).
4. Mortimer, C. H. *J. Ecol.* **29**, 280–329 (1941).
5. Mortimer, C. H. *J. Ecol.* **30**, 147–201 (1942).
6. Cole, G. A. *Textbook of Limnology* (Mosby, St Louis, 1975).
7. Wetzel, R. G. *Limnology* (Saunders, London, 1975).
8. Golterman, H. L. *Physiological Limnology* (Elsevier, Amsterdam, 1975).
9. Delfino, J. J. & Lee, G. F. *Envir. Sci. Technol.* **2**, 1094–1100 (1968).
10. Howard, H. H. & Chisholm, S. W. *Am. midl. Nat.* **93**, 188–197 (1975).
11. Mackereth, F. J. H. *Phil. Trans. R. Soc. B250*, 165–213 (1966).
12. Cunningham, C. R. & Davison, W. *Freshwat. Biol.* **10**, 413–418 (1980).
13. Davison, W. *Limnol. Oceanogr.* **22**, 746–753 (1977).
14. Davison, W. & Rigg, E. *Analyst* **101**, 634–638 (1976).
15. Mackereth, F. J. H., Heron, J. & Talling, J. F. *Scient. Publ. Freshwat. Biol. Ass.* **36**, (1978).
16. Mackereth, F. J. H. *J. scient. Instrum.* **41**, 38–41 (1964).
17. Hesslein, R. H. *Can. J. Fish. aquat. Sci.* **37**, 552–558 (1980).
18. Holdren, G. R. & Bricker, O. P. *ERDA Symp. Ser.* **42**, 178–196 (1977).
19. Morgan, J. J. in *Principles and Applications in Water Chemistry* (eds Faust, S. D. & Hunter, J. V.) 561–624 (Wiley, New York, 1967).
20. Eaton, A. *Geochim. cosmochim. Acta* **43**, 429–432 (1979).
21. Knox, S. & Turner, D. R. *Estuar. coast mar. Sci.* **10**, 317–324 (1980).
22. Duinker, J. C., Wollast, R. & Billen, G. *Estuar. coast mar. Sci.* **9**, 727–738 (1979).
23. Wollast, R., Billen, G. & Duinker, J. C. *Estuar. coast mar. Sci.* **9**, 161–169 (1979).
24. Spencer, D. W. & Brewer, P. G. *J. geophys. Res.* **76**, 5877–5892 (1971).
25. Ruttner, F. *Arch. Hydrobiol. Suppl.* **8**, 197–454 (1931).
26. Verdouw, H. & Dekkers, E. M. J. *Arch. Hydrobiol.* **89**, 509–532 (1980).
27. Dyrssen, D. & Hallberg, R. *Chem. Geol.* **24**, 151–159 (1979).
28. Lu, J. C. S. & Chen, K. Y. in *Chemistry of Marine Sediments* (ed. Yen, T. F.) 181–221 (Ann Arbor, Michigan, 1977).
29. Murray, J. W. & Gill, G. *Geochim. cosmochim. Acta* **42**, 9–19 (1978).

Methylmercury production in the marine water column

G. Topping & I. M. Davies

Department of Agriculture and Fisheries for Scotland, Marine Laboratory, Victoria Road, Aberdeen AB9 8DB, UK

Although the biosynthesis of methylmercury in sediments is well established¹, this is not necessarily the exclusive natural source of methylmercury entering the marine food chain, particularly commercial fish and shellfish species for human consumption. An examination of mercury levels in freshwater fish², collected from a lake with a history of industrial mercury contamination, suggested that levels in fish are controlled in part by mercury in suspension and it followed that methylation should occur in the water column. Although methylmercury is present in seawater in coastal areas receiving discharges of waste containing either inorganic mercury³ or methylmercury⁴ there is no evidence that methylmercury is actually formed in the water column. We now present data which demonstrate that inorganic mercury can be methylated in the water column and we compare this production with that known to occur in marine sediments.

In May 1979 experiments to investigate the movement of inorganic mercury between the water column and sediments were carried out using a large plastic bag (3 m diameter $\times$ 20 m deep) moored in an arm of a sea loch (Loch Ewe) on the west

coast of Scotland. This type of enclosure, which is closed at the bottom, has been used⁵ to study the effects of low levels ($\sim 1 \mu\text{g dm}^{-3}$) of pollutants, particularly metals, on natural populations of phytoplankton, zooplankton and other trophic levels captured in the bags.

In our study, the seawater in the enclosure was initially spiked with inorganic mercury as HgCl_2 , to a concentration of $\sim 0.5 \mu\text{g dm}^{-3}$ using a method described elsewhere⁶. Seawater samples were analysed for total and 'reactive' mercury⁷. Sub-samples of the total settlement material, which was pumped daily from the bottom of the bag, were analysed for total mercury⁸, methylmercury⁹ and carbon (using a CHN analyser). Moderately active primary production was maintained by regular additions of nutrients to the plastic enclosure giving a

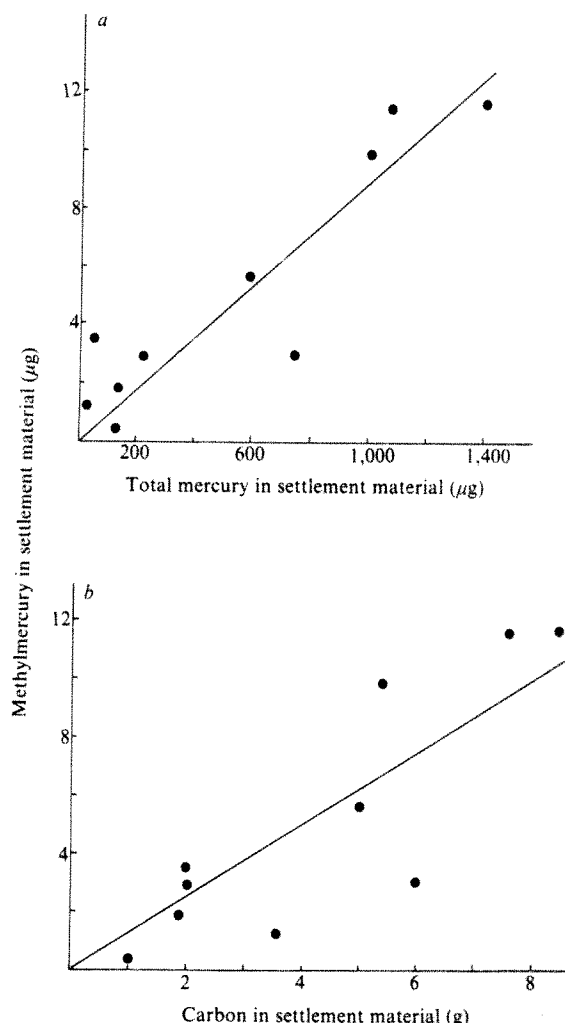


Fig. 1 a, Relationship between methylmercury and total mercury in settlement material. (Regression analysis: $y = 0.00885x$; $r = 0.91$ and standard error of slope = 0.00126.) b, Relationship between methylmercury and carbon in settlement material. (Regression analysis: $y = 1.233x$; $r = 0.85$ and standard error of slope = 0.294.) Both correlations are significant at the 0.01% level.

mean chlorophyll level of 10.9 mg m^{-3} and a mean fallout of organic carbon of $0.4 \text{ g C m}^{-2} \text{ day}^{-1}$.

Significant quantities of methylmercury ($1\text{--}10 \mu\text{g day}^{-1}$) were found in the settlement material. These concentrations were strongly correlated with concentrations of organic carbon ($r = 0.85$) and total mercury ($r = 0.91$) in settlement material; both correlations are significant at the 0.01% level (Fig. 1). Because the seawater in the bag was isolated from the natural sediments in the sea loch throughout the experiment and because the settlement material was not allowed to accumulate for a long

period (about 24 h), we believe that a large part of the methylmercury was formed within the water column.

The amount of methylmercury formed in suspension and collected with the settlement material is directly related to the total mercury associated with the particulate organic matter which in turn is dependant on the level of primary production. From Fig. 1 the ratio methylmercury (μg)/carbon (g) is 1.23 and the ratio methylmercury (μg)/total mercury (μg) is 0.0088; the latter may be compared with 0.001–0.003 estimated by Davies *et al.*³ for a contaminated estuary.

The present experimental results can be used to obtain a rough estimate of the importance of the formation of methylmercury in the water column. If the ratio of methylmercury to mercury of 0.0088 is a guide to what happens in the oceans, and assuming a mean concentration of mercury in the oceans¹⁰ of 10 ng dm^{-3} , the total quantity of methylmercury in open ocean water is 121,000 tonnes. From the global production of carbon¹, our measured ratio of methylmercury (μg) to carbon (g), and assuming that the rate of production of methylmercury is proportional to the total mercury concentration in open ocean, we estimate the amount of methylmercury formed each year in association with carbon fixation at 492 tonnes.

The estimation of methylmercury concentration and production in marine sediments is complicated by the wide range of estimates^{11–16} of methylmercury production rates, from 0.1 – $240 \mu\text{g cm}^{-2} \text{ yr}^{-1}$. Of the two estimates^{14,15} referring to sediments having low mercury concentrations, one¹⁵ includes both water column and sediment production and the other¹⁴ gives a minimum rate of production of $0.56 \mu\text{g cm}^{-2} \text{ yr}^{-1}$ based on uptake by plants and animals in the sediments (containing $2 \mu\text{g cm}^{-3}$ of total mercury). We use the latter value here. We take the mean concentration of mercury in marine sediments¹ as $0.3 \mu\text{g per g wet weight}$ ($0.45 \mu\text{g cm}^{-3}$) and the fraction of mercury present as methylmercury in sediments¹⁶ as 0.1%. If the relevant volume of the sediment compartment is $3.61 \times 10^{19} \text{ cm}^3$ (corresponding roughly to a 10-cm depth over a sea bottom surface of $361 \times 10^{12} \text{ m}^2$) the total methylmercury in marine sediments is 16,200 tonnes. Similarly, production of methylmercury by sediments amounts to 450 tonnes yr^{-1} .

Knauer and Martin¹⁷ found concentrations of 1.3–4.5 ng per g of methylmercury in selected phytoplankton samples from Monterey Bay, California. Using these data (mean value of 3 ng per g) and an annual primary production of $2 \times 10^{17} \text{ g wet weight}$ we estimate that the total amount of methylmercury associated with the annual global production of phytoplankton is ~600 tonnes. This value is comparable with one we obtained for phytoplankton using data from our bag investigations. By comparison, the quantity of methylmercury (~48 tonnes) associated with the annual global production of fish¹ is small.

These data indicate that not only is the 'standing stock' of methylmercury in the oceans greater than that in the surface sediments but that the annual production of methylmercury in these two compartments is similar. Investigations in the field are now necessary to verify our experimental findings and speculations.

Received 15 October 1980; accepted 7 January 1981.

1. *An Assessment of Mercury in the Environment* (National Research Council, National Academy of Sciences, Washington DC, 1978).
2. Armstrong, F. A. J. & Scott, D. P. *J. Fish. Res. Board Can.* **36**, 670 (1979).
3. Davies, I. M., Graham, W. C. & Pirie, J. M. *Mar. Chem.* **7**, 111 (1979).
4. Egawa, H. & Tajima, S. in *Proc. 2nd US Japan Experts Meet.*, Tokyo (1976).
5. *Bull. mar. Sci.* **27**, No 1 (1977).
6. Topping, G. & Windon, H. L. *Bull. mar. Sci.* **27**, 135 (1977).
7. Topping, G. & Pirie, J. M. *Analyt. chim. Acta* **62**, 200 (1972).
8. Topping, G. *Proc. Analyt. Div. chem. Soc.* **14**, 222 (1977).
9. Davies, I. M. *Analyt. chim. Acta* **102**, 189 (1978).
10. Baker, C. W. *Nature* **270**, 230 (1977).
11. Langley, D. G. *J. Wat. Pollut. Cont. Fedn* **45**, 44 (1973).
12. Olsen, B. H. & Cooper, R. C. *Wat. Res.* **10**, 113 (1976).
13. Shin, E. & Krenkel, P. A. *J. Wat. Pollut. Cont. Fedn* **48**, 473 (1976).
14. Windom, H. L., Gardner, W., Stephens, J. & Taylor, F. *Estuarine Coastal mar. Sci.* **4**, 579 (1976).
15. Davies, I. M., Graham, W. C. & Pirie, J. M. *ICES CM* (Fisheries Improvement Committee, E: 31, 1978).
16. Bartlett, P. D., Craig, P. J. & Morton, S. F. *Sci. Total Envir.* **10**, 245 (1978).
17. Knauer, G. A. & Martin, J. H. *Limnol. Oceanogr.* **17**, 868 (1972).

Chondro-osseous morphology of *Dermochelys coriacea*, a marine reptile with mammalian skeletal features

Anders G. J. Rhodin*†, John A. Ogden* & Gerald J. Conlogue*

* Department of Surgery (Orthopaedics), Skeletal Growth and Development Study Unit, Marine Mammal Stranding and Study Center and the Peabody Museum, Yale University, New Haven, Connecticut 06510, USA

† Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA

The marine leatherback turtle, *Dermochelys coriacea* (Testudines, Dermochelyidae), is the largest known extant reptile, reaching weights >680 kg. It has a cosmopolitan oceanic distribution, nesting in tropical climates but migrating widely into colder temperate and boreal waters¹. Recent studies have shown that the leatherback has certain well developed thermoregulatory adaptations: vascular counter-current heat exchangers in the flippers, thick subcutaneous insulation beneath its leathery shell and effective inertial homoiothermy, if not endothermy^{2,3}. Whether this ability to maintain a body temperature higher than the ambient is due to endogenous endothermy or thermal inertia remains unclear^{4,5}. The skeleton of *Dermochelys* remains extensively cartilaginous even in adult animals^{6,7}, which has been attributed to neoteny secondary to its highly pelagic habitat, based on the assumption that its skeleton resembled that of typical embryonic turtles⁸. However, up to now there have been no studies of the internal architecture of the leatherback appendicular skeleton. We report here that the chondro-osseous morphology of *Dermochelys* is unlike that of any other known extant turtle or reptile but is more similar to that of marine mammals, notably Cetacea (whales) and Sirenia (manatees).

The typical hard-shelled marine turtles (Cheloniidae—genera *Chelonia*, *Caretta*, *Eretmochelys* and *Lepidochelys*) show a pattern of chondro-osseous morphology highly reminiscent of terrestrial or aquatic turtles^{9,10}. A thick, distinct compacta is clearly delineated from a medullary cancellous region and the epiphyses are covered by a very thin avascular cartilage which serves the purpose both of articular cartilage and physis, or growth plate. The subchondral plates of the epiphyseal bone surfaces are smooth and perfectly parallel to the articular cartilages. The only difference in marine turtles is the failure of development of a medullary cavity.

Dermochelys shows several important differences in skeletal morphology. In dried whole bones (with cartilage removed) the epiphyseal subchondral plates are not smooth, but instead present a rough, undulated, fenestrated surface which is not parallel to the contours of the articular cartilage surface (Fig. 1). In longitudinal sections of fixed whole bones each epiphyseal end has large cartilaginous epiphyses. Each of these, whether large or small, is filled with an extensive cartilage canal vascular system (Fig. 1). The vessels are both perichondral and transphyseal in origin and allow continuity of circulation between metaphysis and epiphysis. Both large and small vessels cross the physis: the smaller ones probably participate in endochondral chondro-osseous replacement, and the large ones are probably involved in cartilaginous expansion and nutrition of the epiphysis itself. No secondary calcification or ossification centres develop in any of these cartilaginous areas.

The diaphysis and metaphysis are filled with relatively dense cancellous trabecular bone without secondary medullary cavity formation. The metaphysis seems to be almost totally derived from endochondral bone formation and the diaphysis from a combination of endochondral and periosteal membranous growth. This is dramatically evident because of a combination of

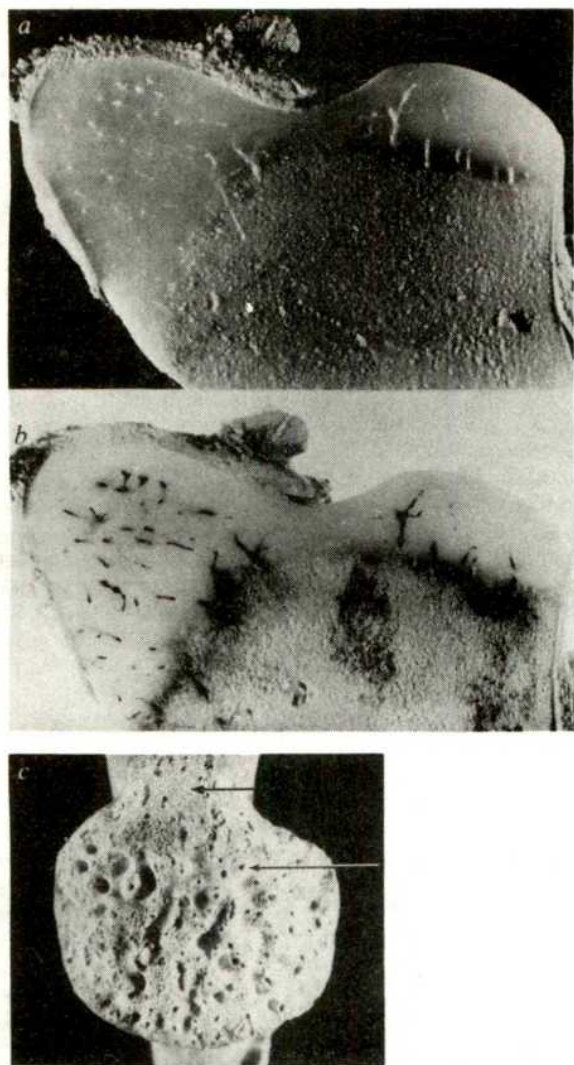


Fig. 1 Dried whole humeri were prepared by dermestid beetle maceration from two *Dermochelys coriacea*, one *Eretmochelys imbricata* (hawksbill), three *Chelonia mydas* (green turtle), two *Caretta caretta* (loggerhead) and two *Lepidochelys kempii* (ridley). Sections of whole formalin-fixed humeri, radii, ulnas and metacarpals were examined from two *C. caretta*, two *L. kempii* and four *D. coriacea* (straight-line carapace lengths 135, 153, 175 and 178 cm, estimated weights 260, 360, 470 and 470 kg; specimens found stranded in New Jersey during August–October 1979, and obtained from the Marine Mammal Stranding Center, Atlantic City). *a*, *b*, Natural and dye-injected longitudinal sections of the proximal humerus in a *D. coriacea* of 135-cm carapace length: medial process apophysis on the left, proximal epiphysis on the right. Small vessels of perichondral origin are seen in the centre of the apophysis; large transphyseal vessels are seen in both the epiphysis and the apophysis. *c*, View end-on of dried proximal humeral epiphyseal subchondral plate (cartilage removed) in a *D. coriacea* of ~190-cm carapace length. Arrowhead points to small metaphyseal end-arterioles participating in endochondral osteogenesis; black arrow points to large transphyseal fenestrations allowing vascular access to the centre of the cartilaginous epiphysis. Also seen are coarse growth plate undulations, which probably serve as stabilization for the large chondroepiphysis.

dark pigmentation in the periosteally derived trabecular bone and lack of internal ontogenetic remodelling (Fig. 2). Cones of light-coloured endochondral and dark-coloured periosteal bone are therefore formed, which radiate in a well delineated manner from the central nutrient artery located at the anatomic site of the initial diaphyseal ossification centre (see radiograph, Fig. 2).

Transverse sections of the mid-humerus show a cyclical variation in bone deposition which is mainly peripheral, both in the endochondral and periosteal portions. Other than pigmentation differences it is difficult to delineate grossly endochondral and

periosteal bone and the impression gained is of an amedullary cancellous region merging gradually peripherally into a slightly denser compacta, which is extensively vascular in the deeper portions but less so superficially, where it becomes a very thin shell of dense subperiosteal bone.

In reconstructing the hypothetical growth of this bone, it seems that both patterns of bone formation occur contiguously for a certain period of time, with the endochondral cones of bone progressively diverging from each other at their bases and the intervening space being filled in by periosteal bone. However, in the sub-adult animal it seems that further elongation of the bone becomes solely derived from endochondral growth, with very little contribution from the periosteal bone apart from the continued formation of a very thin cortical compacta. This apparent change in growth pattern may represent a relative onset of skeletal maturity comparable with closure of the physis in mammals. However, as secondary epiphyseal ossification and growth-plate fusion do not occur in the leatherback, it is capable of continual growth throughout adulthood, this apparently being mainly endochondral.

In all these features of chondro-osseous morphology the leatherback turtle appears most similar to marine mammals, notably the Cetacea (whales)¹¹ and Sirenia (manatees)¹². Both of these groups also vascularize their epiphyses perichondrally and transphyseally, develop amedullary bones with retained endochondral and periosteal cones that are pigmented and do not remodel, show a gradual merging of cancellous into compact bone, and in some forms (the migratory beluga whale) also show cyclical bone deposition^{11,12}. The only discernible differences in the leatherback are the lack of secondary epiphyseal calcification and ossification, the point of relative cessation of periosteal growth in the sub-adult animal, and the presence of dark pigmentation in the periosteal rather than the endochondral bone, which is a reversal of the pattern seen in the marine mammals.

No other known extant reptile shows this combination of chondro-osseous developmental features. Whereas some extant reptiles (for example, the varanid lizards, including the large Komodo dragon) vascularize their chondroepiphyses, this vascularization is always perichondrally or circumphyseally derived, never transphyseally¹³. In addition, the vascularization is invariably followed by secondary calcification and ossification of the epiphysis. All other turtles and all crocodilians previously examined, including species of large body size, fail to vascularize their thin chondroepiphyses, which also remain cartilaginous throughout adulthood^{8,13}. However, certain extinct reptiles show some similarities to the leatherback: plesiosaurs have been described as having endochondral and periosteal cones that do not remodel¹⁴; ichthyosaurs, plesiosaurs, nothosaurs and mesosaurs are known to have amedullary bones with cancellous-compacta differentiation paralleling that of marine mammals^{15,16}, and protostegid turtles have also been noted briefly to be somewhat similar in this respect¹⁷. All these fossil reptiles were highly adapted to a marine existence as shown by other skeletal features. Some fossil amphibians with pronounced marine modifications (certain stegocephalian species) also show a pattern of bone histology very similar to that described for marine mammals¹⁸. Other extant vertebrates with marine adaptations have similar developmental patterns, notably pinnipeds (seals)¹⁹ and penguins²⁰. However, some marine vertebrates do not show this spectrum of characteristics—notably the typical hard-shelled marine turtles (Cheloniidae) which are, in most respects, similar to freshwater or terrestrial species.

No terrestrial vertebrate has these chondro-osseous developmental features. The fact that such diverse groups as cetaceans, sirenians, pinnipeds, penguins, extinct marine reptiles and amphibians, and leatherback turtles have such a high degree of physical similarity in bone morphology suggests an underlying mechanism of marine adaptability which has led to a highly developed pattern of skeletal evolutionary convergence. The failure of some marine vertebrates (for example, hard-shelled turtles) to develop these characteristics fully may

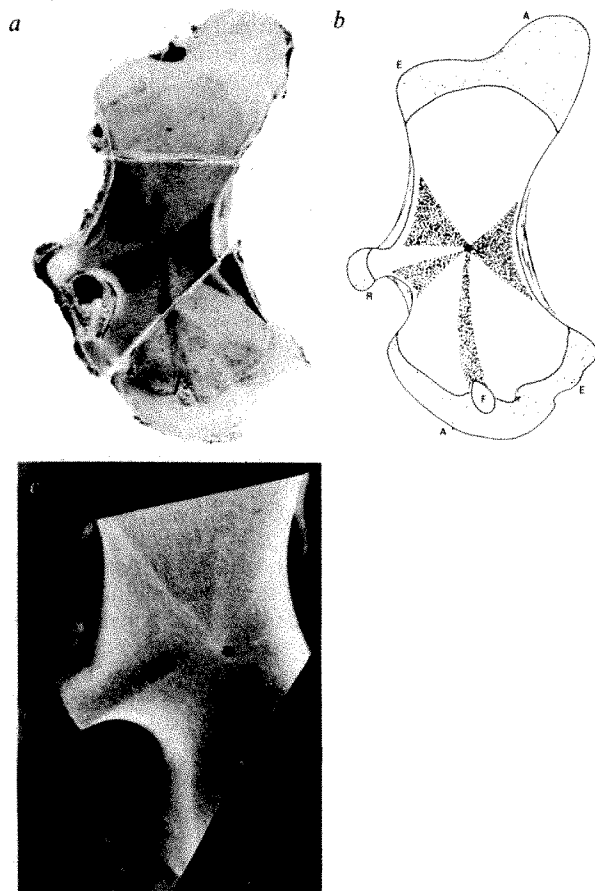


Fig. 2 a, b, Actual and schematic longitudinal sections of unstained humerus of a *D. coriacea* of 135-cm carapace length. Endochondral bone is light coloured whereas membranous (periosteal) bone is dark. Note also the amedullarity, lack of clear cancellous-compacta differentiation and disappearance of pigmentation at the superficial cortical margin. E, epiphyses; A, apophyses; F, ectepicondylar foramen; R, radial process apophysis. c, Radiograph of thin longitudinal slice of central diaphyseal portion of humerus showing central nutrient foramen and radiating cones of endochondral and membranous bone. Orientation as in Fig. 2a.

be related either to an incompletely evolved system due to relatively recent marine adaptation, or to a different set of physiological demands being required for differing life habits. Some studies have suggested that the marine pattern of bone development may be related to the homeostatic requirements of prolonged deep diving^{19,21,22}. It has been hypothesized that the intense dysbaric and acidotic challenges produced by repetitive deep diving can be countered by a skeletal buffering system which becomes more effective through increased vascularity in denser, amedullary bone¹⁹. Probably also of significance are the altered skeletal biomechanics of a marine existence where the limbs are not needed to support weight but instead are subject to entirely different forces produced by propulsion or steering in a medium much denser than air. Theories have also been advanced supporting hypothyroidism¹² and altered thermal relationships¹¹ as the underlying mechanisms leading to marine skeletal characteristics.

With regard to the phylogeny of *Dermochelys*, theories have vacillated between those supporting extreme specialization in a recently derived form and those favouring extreme primitiveness. Current taxonomic opinion accepts the former theory and is based on many similarities to the hard-shelled turtles²³⁻²⁸. However, strong arguments supporting the latter have also been advanced in the past^{7,8,29,30}. We suggest that the chondro-osseous morphology of *Dermochelys* emphasizes its extreme distinctness from other extant marine turtles (and turtles and reptiles in general). Whether this distinctness is recently derived or primitive cannot yet be answered.

Theories concerned with the phylogeny of endothermy and the origin of mammals from reptiles have stressed the relationships of increased body size and thermal inertia³¹ as well as altered bone histology³². It has been suggested that endothermy evolved separately in a number of extinct reptile groups³¹, and that transphyseal vascularization of a permanently cartilaginous epiphysis may have been the pattern in early mammal-like reptiles³³. The striking similarities in bone growth between *Dermochelys* and certain marine mammals, the presence of transphyseal vessels without secondary epiphyseal ossification, and the well developed homoiothermic (or endothermic) thermoregulatory adaptations suggest that this large animal may have developed along a course parallel to certain extinct endothermic reptiles, although in a group not previously thought to be undergoing this evolutionary process. In addition, the leatherback turtle's chondro-osseous morphology represents a striking and seemingly unique parallel between an extant reptile and modern marine mammals.

We thank the following people for their help and comments: R. C. Schoelkopf, A. Mead, J. P. Ross, E. E. Williams, H. C. W. Skinner and C. R. Shoop. This work was supported in part by grants R01-HD-10854 and K04-AM-00300 from the NIH.

Received 18 September 1980; accepted 21 January 1981.

1. Pritchard, P. C. H. *Catalog of American Amphibians and Reptiles*, 238 (1980).
2. Frair, W., Ackman, R. G. & Mrosovsky, N. *Science* **177**, 791 (1972).
3. Greer, A. E., Lazell, J. D. Jr & Wright, R. M. *Nature* **244**, 181 (1973).
4. Neill, W. H. *et al. Science* **184**, 1008 (1974).
5. McNab, B. K. & Auffenberg, W. *Comp. Biochem. Physiol.* **55A**, 345 (1976).
6. Gervais, P. *Mém. Nouv. Arch. Mus. nat. Hist., Paris* **8**, 199 (1872).
7. Dollo, L. *Bull. Mus. R. Hist. nat. Belg.* **5**, 59 (1888).
8. Deraniyagala, P. E. P. *Proc. zool. Soc. Lond.* **68**, 1057 (1930).
9. Haines, R. W. *J. Anat.* **72**, 323 (1938).
10. Suzuki, H. K. *Ann. N.Y. Acad. Sci.* **109**, 351 (1963).
11. Felts, W. J. L. & Spurrell, F. A. *Am. J. Anat.* **118**, 103 (1966).
12. Fawcett, D. W. *Am. J. Anat.* **71**, 271 (1942).
13. Haines, R. W. *J. Anat.* **75**, 282 (1941); *Biol. Rev.* **17**, 267 (1942); in *Biology of the Reptilia* Vol. 1 (eds Gans, C., Bellairs, A. d'A. & Parsons, T. S.) 81-115 (Academic, London, 1969).
14. Moodie, R. L. *Am. J. Anat.* **7**, 443 (1908).
15. Enlow, D. H. & Brown, S. O. *Tex. J. Sci.* **9**, 186 (1957).
16. Ricqlès, A. J. de in *Morphology and Biology of Reptiles* (eds Bellairs, A. d'A. & Cox, C. B.) 123-150 (Academic, London, 1976).
17. Cope, E. D. *Proc. Am. phil. Soc.* **12**, 422 (1872).
18. Ricqlès, A. J. de *Colloques int. Cent. nat. Rech. Sci.* **218**, 351 (1975).
19. Versaggi, C. S. *Sci. Biol. J.* **3**, 305 (1977).
20. Meister, W. *Anat. Rec.* **143**, 377 (1962).
21. Nopcsa, F. *Anat. Anz.* **56**, 353 (1923).
22. Sickenberg, O. *Palaeobiologica* **4**, 405 (1931).
23. Case, E. C. *J. Morph.* **14**, 21 (1897).
24. Wieland, G. R. *Am. J. Sci.* **9**, 413 (1900); **14**, 95 (1902); **27**, 101 (1909).
25. Romer, A. S. *Osteology of the Reptiles* (University of Chicago Press, Chicago, 1956).
26. Gaffney, E. S. *Bull. Am. Mus. nat. Hist.* **155**, 387 (1975).
27. Wermuth, H. & Mertens, R. *Tierreich* **100**, 1 (1977).
28. Frair, W. *Herpetologica* **35**, 239 (1979).
29. Hay, O. P. *Am. Nat.* **32**, 929 (1898).
30. Zangerl, R. *Fieldiana, Geol. Mem.* **3**, 57 (1953).
31. McNab, B. K. *Am. Nat.* **112**, 1 (1978).
32. Ricqlès, A. J. de *C. r. hebdom. Séanc. Acad. Sci., Paris* **268D**, 782 (1969); *Evol. Theory* **1**, 51 (1974); in *Morphology and Biology of Reptiles* (eds Bellairs, A. d'A. & Cox, C. B.) 123-150 (Academic, London, 1976).
33. Ricqlès, A. J. de *Année biol.* **18**, 1 (1979).

Genetic control for sibling recognition?

Andrew R. Blaustein & Richard K. O'Hara

Department of Zoology, Oregon State University, Corvallis, Oregon 97331, USA

The ability to distinguish between kin and non-kin is critical to current theories of altruistic behaviour and kin selection¹. Hamilton¹ predicted that individuals would behave differently towards one another depending on the genetic relatedness between them. When either proximity to or familiarity with kin is a good predictor of relatedness, the mechanism by which favouritism towards kin is accomplished may not require special kin recognition abilities^{2,3}. However, if proximity and familiarity are poor predictors of kinship, favouritism (hence increases in inclusive fitness⁴) could only be achieved by the differential

recognition of kin and non-kin. We have previously shown that *Cascades frog (Rana cascadae)* tadpoles reared with siblings prefer to associate with siblings over non-siblings⁵. The present study is the first to report that totally naive individuals (*R. cascadae* tadpoles) prefer to associate with siblings over non-siblings. Because tadpoles were separated before hatching and reared apart from other individuals, results suggest that the ability of these tadpoles to discriminate between siblings and non-siblings has some innate component.

Sibling recognition has been reported in experimental studies of spiny mice (*Acomys cahirinus*)^{6,7}, pigtail macaques (*Macaca nemestrina*)⁸ and American toad tadpoles (*Bufo americanus*)⁹. In each case, young were separated after birth or hatching and later given a choice of associating with siblings and non-siblings. In the mouse studies, sibling recognition was probably a result of learning during early development^{6,7}. Sibling recognition in pigtail macaques is possible in the absence of prior sibling association. It is unclear how sibling recognition develops in toad tadpoles. Although the existence of a genetic recognition system has been proposed in theoretical discussions of kin selection^{1,10,11}, experimental evidence (with the possible exception of the macaque study⁸) is lacking.

Two *R. cascadae* clutches were collected in the Oregon Cascade Mountains and transported to the laboratory in separate containers. Twenty eggs from one clutch (group A) were placed individually in opaque containers (A_1 individuals) at early embryonic development (yolk plug stage; stage 12, ref. 12). The remainder of the clutch (250 tadpoles) was reared in a single aquarium (group A). A second clutch was divided in half (groups B_1 and B_2 ; developmental stage 19, ref. 12) and each group was placed in a separate aquarium.

A rectangular tank (122 × 44 × 30 cm) was used to test two tadpoles simultaneously for sib group preferences. To create end compartments, a partition of 1.5-mm plastic mesh was placed 15 cm from each end of the tank. The remaining central portion of the tank was longitudinally divided by an opaque, water-tight partition. When testing A_1 individuals, 25 tadpoles from group A and 25 tadpoles from group B_1 were placed in opposite ends of the tank. When testing B_1 individuals, 25 tadpoles from groups A and B_2 were placed in opposite ends of the tank. For each test, these stimulus tadpoles were left undisturbed for 15 min. Then, at the tank centre, one tadpole was released on each side of the partition and allowed to acclimatize for 10 min. Each tadpole was tested in four 5-min trials at 10-min intervals. We recorded time spent in seconds in sib and non-sib halves of the tank.

Tadpoles reared in total isolation since early embryonic development (A_1) and tadpoles reared with sibs (B_1) but given a choice of unfamiliar sibs or non-sibs preferred to associate with sibs. Both the number of individuals spending most of their time on the sib side of the tank (18/20 for A_1 and 16/20 for B_1 ; $P = 0.001$ and $P = 0.006$ respectively, binomial test) and the number of seconds spent by each of these individuals were significant (Table 1). These results provide the first experimental evidence that an organism can recognize siblings after being separated from sibs while still embryos and after being raised in complete isolation.

Long-term field observations on the social and ecological conditions surrounding early development of *R. cascadae* indicate that ample opportunity exists for sibling groups to develop¹³. Females oviposit eggs in specific habitats in one or more localities around a lake or pond. Clutches are usually laid next to or on top of other conspecific egg masses. As they develop, dispersal from the site of oviposition tends to be low. Furthermore, larvae were never observed singly but always in close association with conspecifics. This tendency for *R. cascadae* to aggregate has been verified experimentally¹³. It is likely that natural groupings of this species are comprised of a high proportion of siblings and early development occurs in the presence of sibs and non-sibs.

Predation is probably one selective force for group formation in *R. cascadae* as they are palatable and preyed on by several

Table 1 Mean times out of 1,200 s spent by individuals on sib side of test tank when given a choice between sibs and non-sibs

Individuals tested	n	Seconds spent on sib side (mean ± s.e.m.)	T
A_1	20	702.6 ± 20.65	15*
B_1	20	682.2 ± 33.22	50†

Tests of A_1 tadpoles and B_1 tadpoles were made at 14–18 and 19–23 days post-hatching respectively. The Wilcoxon signed rank test (one-tailed; t represents the test statistic)¹⁸ was used to test for significant departures from an expected random distribution of total times on sib and non-sib ends. Individuals from group A_1 were reared in 0.5-l containers. The other members of groups A, B_1 and B_2 were reared in 38-l aquaria. All tadpoles were reared in the same room in a 14-h light 10-h dark photoperiod. The temperature in the laboratory was kept at 20–22 °C. Tadpoles were fed rabbit pellets *ad libitum*. The same 25 stimulus individuals were used for five tests, and were then discarded. Between tests, end compartment placements of As and Bs were alternated. Previously, two controls (20 tadpoles each from two different clutches) were run in the absence of stimulus groups to determine possible biases in the testing apparatus. In these controls tadpoles displayed a random distribution pattern⁵. Following each test, tadpoles were removed, the tank drained, rinsed and refilled to a depth of 6 cm with 34 l of dechlorinated tap water. No tadpole was tested more than once. Average developmental stages¹² for groups A_1 , A and B were 28.5, 28.0 and 28.0 respectively.

* $P < 0.005$; † $P < 0.025$.

species^{13,14}. Individuals can increase their fitness by living in groups and detection of predators increases with increasing group size¹⁵. Although kinship alone is not a necessary prerequisite for the evolution of groups, an individual in a group consisting of many sibs who warns others of a predator may increase its inclusive as well as its individual fitness compared with an individual giving a similar warning in a group with fewer or no sibs^{16,17}.

Our data may be added to the few reports^{6–9} which suggest that some species preferentially associate with sibs over non-sibs. Significantly, there is a major similarity and a major difference between the present study and that of Wu *et al.*⁸ which greatly enhance the conclusions of both studies. The use of paternal half sibs by Wu *et al.*⁸ and our use of individuals reared in isolation from an early embryonic stage show that individuals can recognize sibs without prior sib association. However, the macaques used by Wu *et al.*⁸ were allowed to associate with non-sib conspecifics during development whereas tadpoles in our study were reared in total isolation and had no exposure to conspecifics during development. Furthermore, we suggest that the ability to recognize sibs in *R. cascadae* may exist in early embryonic life. The results of this and another study to be reported in detail elsewhere⁵ are consistent with a genetic recognition system. In the latter study, two different clutches were reared from an embryonic stage in an aquarium separated by plastic mesh. This allowed chemical and visual contact between tadpoles during development. When given a choice of unfamiliar sibs or unfamiliar non-sibs, these tadpoles associated with unfamiliar sibs. When given a choice between a group comprised of an equal mixture of sibs and non-sibs (tadpoles with which they were reared) and an unfamiliar sib group, tadpoles associated with unfamiliar sibs. Although it is possible that individually reared tadpoles could imprint on their own cues (or subtle maternal cues from the egg mass jelly), results of tadpoles reared with non-sibs⁵ argue against this.

In organisms that undergo early development in the presence of non-kin (such as *R. cascadae*) or in those that disperse before associating with kin, a kin recognition mechanism that functions independently of experience may be required to achieve recognition. Our experiments are consistent with such a mechanism.

We thank P. S. Dawson, W. G. Holmes, R. H. Porter, W. R. Rice, P. B. Samollow, R. R. Warner and J. A. Wiens for valuable comments. M. Bekoff, R. H. Porter and W. R. Rice critically reviewed the manuscript.

Received 3 October 1980; accepted 13 January 1981.

1. Hamilton, W. D. *J. theor. Biol.* **7**, 1–52 (1964).
2. Sherman, P. W. in *Sociobiology: Beyond Nature/Nurture* (eds Barlow, G. W. & Silverberg, J.) 505–544 (Westview, Boulder, 1980).
3. Bekoff, M. in *Parental Care in Mammals* (eds Gubernick, D. & Klopfer, P. H.) 307–346 (Plenum, New York, 1981).
4. Wilson, E. O. *Sociobiology: The New Synthesis* (Harvard University Press, 1975).
5. O'Hara, R. K. & Blaustein, A. R. *Anim. Behav.* (submitted).
6. Porter, R. H., Wyrick, M. & Pankey, J. *Behav. Ecol. Sociobiol.* **3**, 61–68 (1978).
7. Porter, R. H. & Wyrick, M. *Anim. Behav.* **27**, 761–766 (1979).
8. Wu, H. M. H., Holmes, W. G., Medina, S. R. & Sackett, C. P. *Nature* **285**, 225–227 (1980).
9. Waldman, B. & Adler, K. *Nature* **282**, 611–613 (1979).
10. Dawkins, R. *The Selfish Gene* (Oxford University Press, 1976).
11. Rothstein, S. I. *J. theor. Biol.* **87**, 255–261 (1980).
12. Gosner, K. L. *Herpetologica* **16**, 183–190 (1960).
13. O'Hara, R. K. thesis, Oregon State Univ. (1981).
14. Briggs, J. L. & Storm, R. M. *Herpetologica* **26**, 283–300 (1970).
15. Bertram, B. C. R. in *Behavioural Ecology: An Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 64–96 (Sinauer, Sunderland, 1978).
16. Sherman, P. W. *Am. Nat.* **113**, 925–935 (1979).
17. West-Eberhard, M. J. *Q. Rev. Biol.* **50**, 1–34 (1975).
18. Hollander, M. & Wolfe, D. A. *Nonparametric Statistical Methods* (Wiley, New York, 1973).

Substrate-adhering lymphoid cells show impaired tumorigenicity and increased immunogenicity

Jacob Hochman*, Aviva Katz, Efrat Levy & Shmuel Eshel

Department of Zoology, The Hebrew University of Jerusalem, Jerusalem, Israel

Cell adhesiveness is fundamental to a variety of biological phenomena, including tumour development and metastasis^{1–4}. Recently, Bubenik *et al.*⁵ have reported that in various malignant fibroblastoid cell lines those cells which demonstrated less adhesiveness were more tumorigenic. The relationship between cell adhesiveness, transformation and metastasis has been studied extensively in cells (fibroblastoid) grown as monolayers attached to their substratum and to each other^{4,6–9}, but, to our knowledge, there has been no report describing this relationship in suspension-borne (lymphoid) cells that grow free of each other and their substratum. We report here that substrate-adhering variant cells, selected from the tumorigenic, suspension-growing S49 mouse lymphoma, have impaired tumorigenicity. Furthermore, the substrate-adhering cells also have increased immunogenicity, as their inoculation into mice protects the mice from subsequent challenges with parental, non-adherent tumorigenic S49 cells. These findings suggest a new approach for the selection of immunogenic cells from suspension-borne tumorigenic cell populations.

In our study, a subline of S49 mouse lymphoma cells¹⁰, descendants of clone 24.6.1 (ref. 11), was propagated in suspension culture and readapted to grow as a transplantable tumour in BALB/c mice¹². Selection for cells with increased tumorigenicity was carried out through serial passages as tumours in BALB/c mice. Tumour cells from the 25th passage, which formed tumours as early as 7–10 days post-inoculation, were transferred into culture medium (subline T-25). These cells retained the same tumorigenicity *in vivo* for at least 6 months after transfer into suspension culture. From T-25 cells we have selected without the use of mutagens a subline (designated T-25-Adh) able to adhere to the substrate bottom of tissue culture flasks (Table 1). This selection involved continuous enrichment of substrate-adhering cells by discarding the cells growing in suspension. While attached to their substrate, these cells continued to grow until a dense monolayer was formed.

Table 1 Comparison between parental T-25 cells and their adherent variant (T-25-Adh) subline

Cells	Substrate adhesiveness	Tumorigenicity*	Immunizing† ability against T-25 cells in syngeneic mice
T-25	–	+	–
T-25-Adh	+	–	+

Cells were propagated in culture in Dulbecco's modified Eagle's medium (Gibco), supplemented with 10% heat-inactivated horse serum, penicillin (50 U ml^{–1}) and streptomycin (50 µg ml^{–1}). Cells were maintained at 37 °C in a humidified atmosphere containing 5% CO₂. Adhering cells were detached from their substratum (bottom of tissue culture flask) by shaking. Cells were adjusted to 1 × 10⁶ cells ml^{–1} and 10-ml portions were added to 50-ml tissue culture flasks (Costar 3050, 25 cm²) and incubated at 37 °C. At various intervals thereafter, flasks were tilted, aliquots taken out and the number of cells remaining unattached scored. From this, the percentage of attached cells was calculated. Although none of the parental T-25 cells attached, essentially all T-25-Adh cells attached to the substrate within 45–60 min. Assays were carried out only on samples in which viability, according to the trypan blue exclusion test, was >95%.

* BALB/c males, 8–10 weeks old, were inoculated with 1 × 10⁷ cells intraperitoneally. Mice were checked daily, and the first day on which tumours could be recognized was recorded. T-25 cells (*n* = 16) developed tumours within 7–10 days and all the mice died within 13–15 days. T-25-Adh cells (*n* = 14) did not develop any tumours for at least 90 days.

† BALB/c males inoculated intraperitoneally 90 days earlier with 1 × 10⁷ T-25-Adh cells (*n* = 10) and control BALB/c males (*n* = 4) were both inoculated with tumorigenic T-25 cells (1 × 10⁷ cells per mouse, intraperitoneally). Control mice developed tumours and died within 15 days, whereas all the mice that had previously received T-25-Adh cells did not develop any visible tumours for at least 180 days.

When the substrate was fully covered by a monolayer, cells appeared in the liquid phase of the medium above. These cells then continued to divide normally as a suspension of single cells or small aggregates of a few cells. When the free-floating cells were transferred to a new culture flask, they attached to the substrate with the same kinetics as their parental adherent cells.

When T-25 and T-25-Adh cells were inoculated into BALB/c mice, only mice that had received the T-25 (non-adherent) cells developed tumours (100% within 10 days) and died within 13–15 days post-inoculation. On the other hand, mice that received the same inoculum of T-25-Adh (substrate-adhering) cells did not demonstrate any visible tumours for at least 90 days (Table 1). These findings demonstrate that the selection for substrate-adhering S49 cells significantly decreases their tumorigenicity in BALB/c mice.

We tested whether the decrease in tumorigenicity of the substrate-adhering cells involved an immune response component, by exposing BALB/c mice that had previously received T-25-Adh cells to a challenge with 10⁷ tumorigenic T-25 cells per mouse. BALB/c mice of the same age group that had not been previously subjected to T-25-Adh cells were controls (Table 1). T-25-Adh cells were able to confer on BALB/c mice immunity against a challenge of suspension-growing tumorigenic T-25 cells. Furthermore, in preliminary experiments, the immunizing ability against T-25 cells could be adoptively transferred to BALB/c mice via spleen cells from mice previously inoculated with adherent (T-25-Adh) cells.

Both the selection for substrate adhesiveness and immunization procedures have been repeated with another tumorigenic subline, T-40 (derived from the 40th passage *in vivo*), and its selected adherent variant, and a tumorigenic subline derived from descendants of a different clone (24.3.2) of S49 cells¹¹. Here again, the adherent variant cells demonstrated impaired tumorigenicity concomitantly with increased immunogenicity when compared with their suspension-growing parental cells.

* Present address: Laboratory of Molecular Biology, National Cancer Institute, NIH, Bethesda, Maryland 20205, USA.

This report demonstrates for the first time that the relation between increased cell adhesiveness and decreased tumorigenicity^{5,9,13} may be applied to malignant cells of the lymphoid category. Furthermore, the impaired tumorigenicity of substrate-adherent S49 cells is, at least in part, the result of an increase in their immunizing ability.

Xenogenization or heterogenization are general terms now used to describe all attempts to make tumour cells antigenically foreign to their host so that they may be used more effectively in immunotherapy and immunodiagnosis. Several approaches are used for the biological xenogenization of tumour cells¹⁴⁻²¹, a comprehensive summary of which was published recently²². Our findings in S49 cells suggest a new and somewhat different approach for the xenogenization of malignant lymphoid cells. Here, the isolation of immunogenic, non-tumorigenic cells is based on a simple selection for substrate adherence. The nature of the cell-surface change(s) that took place on transition from free-floating to substrate-adhering, non-tumorigenic, immunogenic S49 cells is being investigated. Furthermore, it is of both mechanistic and practical significance to find out whether the approach taken here is also applicable to other lymphoid tumour cell lines.

This work was supported by a special fund from the faculty of Mathematics and Natural Sciences, The Hebrew University of Jerusalem, to J. H.

Received 26 September 1980; accepted 7 January 1981.

1. Ludford, R. J. *Proc. R. Soc. B* **112**, 250-263 (1932).
2. Cowdry, E. V. *Archs Path.* **30**, 1245-1274 (1940).
3. Coman, D. R. *Cancer Res.* **4**, 625-647 (1944).
4. Grinnell, F. *Int. Rev. Cytol.* **53**, 65-143 (1978).
5. Bubenik, J. et al. *Int. J. Cancer* **23**, 392-395 (1979).
6. Pouyssegur, J. M. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **73**, 544-548 (1976).
7. Winkelhake, J. L. & Nicolson, G. L. *J. natn. Cancer Inst.* **56**, 285-291 (1976).
8. Briles, E. B. & Kornfeld, S. *J. natn. Cancer Inst.* **60**, 1217-1222 (1978).
9. Murray, J. C., Liotta, L., Rennard, S. I. & Martin, G. R. *Cancer Res.* **40**, 347-351 (1980).
10. Horiata, K. & Harris, A. W. *Exp. Cell Res.* **60**, 61-77 (1970).
11. Coffino, P., Bourne, H. R. & Tomkins, G. M. *J. cell. Physiol.* **85**, 603-610 (1975).
12. Hochman, J., Katz, A., Levy, E. & Eshel, S. (in preparation).
13. Coman, D. R. *Cancer Res.* **14**, 519-531 (1954).
14. Lindenmann, J. *J. Immun.* **92**, 912-919 (1964).
15. Lindenmann, J. *Biochim. biophys. Acta* **355**, 49-75 (1974).
16. Bekesi, G., Arneault, S. & Holland, J. *Proc. Am. Ass. Cancer Res.* **12**, 47 (1971).
17. Hashimoto, Y. & Yamanoha, B. *Gann* **67**, 315-319 (1976).
18. Hashimoto, Y., Yamanoha, B., Endoh, H. & Ishizuka, T. in *Gann Monograph on Cancer Research* (ed. Kobayashi, H.) **23**, 135-141 (Japan Scientific Societies, Tokyo, 1979).
19. Harris, H., Miller, O. L., Klein, G., Worst, P. & Tachibana, T. *Nature* **223**, 363-366 (1969).
20. Klein, G. & Harris, H. *Nature new Biol.* **237**, 163-164 (1972).
21. Shinitzky, M., Skornick, Y. & Haran-Ghera, N. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5313-5316 (1979).
22. *Gann Monograph on Cancer Research* (ed. Kobayashi, H.) **23** (Japan Scientific Societies, Tokyo, 1979).

Fibroblast traction as a mechanism for collagen morphogenesis

Albert K. Harris, David Stopak & Patricia Wild

Department of Zoology, University of North Carolina, Wilson Hall 046-A, Chapel Hill, North Carolina 27514, USA

To make visible the traction forces exerted by individual cells, we have previously developed a method of culturing them on thin distortable sheets of silicone rubber¹. We have now used this method to compare the forces exerted by various differentiated cell types and have examined the effects of cellular traction on re-precipitated collagen matrices. We find that the strength of cellular traction differs greatly between cell types and this traction is paradoxically weakest in the most mobile and invasive cells (leukocytes and nerve growth cones). Untransformed fibroblasts exert forces very much larger than those actually needed for locomotion. This strong traction distorts collagen gels dramatically, creating patterns similar to tendons and organ capsules. We propose that this morphogenetic rearrangement of extracellular matrices is the primary function of fibroblast traction and explains its excessive strength.

As a tissue culture cell spreads on one of our silicone rubber substrata, the traction force with which the cell margin pulls itself forwards compresses an area of the rubber sheet beneath the cell (Fig. 1a). The thin rubber sheet becomes folded into wrinkles of two kinds: compression wrinkles directly beneath the cell, running transverse to the force, and tension wrinkles which radiate outwards from the cell margins in the direction of maximum stretching. Of the cells we have examined, the strongest traction was exerted by fibroblasts from embryonic chick heart, skeletal muscle and tendon, dermis and the choroid coat of the eye, and by BALB/3T3 mouse cells. Individual cells of these types produced numerous wrinkles beneath and around themselves (Fig. 1a, b). Glial cells from dorsal root ganglia and platelets from human blood also exerted very strong traction. In contrast, epithelial cells (from embryonic chick epidermis, pigmented retina and liver parenchyma) were considerably weaker. These could still distort the rubber sheet, but only when many

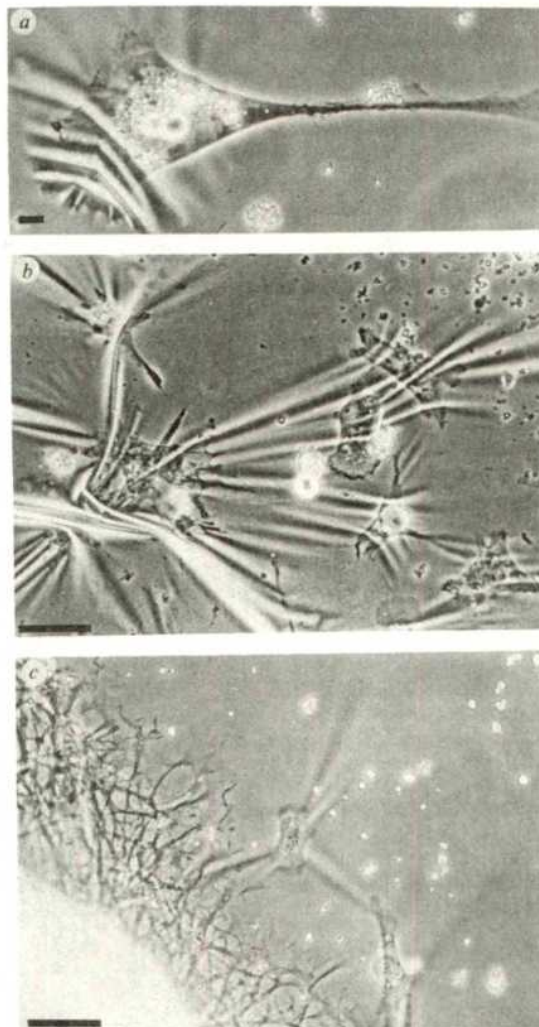


Fig. 1 Tissue culture cell spreading on thin layers of silicone rubber. The traction by which the cells spread and propel themselves is shown by the wrinkles formed in the rubber sheet. Silicone rubber substrata were made by flaming the surface of 30,000 cP silicone fluid (polydimethylsiloxane, Dow-Corning; see ref. 1 for details of method). *a*, An individual 3T3 cell which has compressed the rubber layer beneath its leading margin into numerous folds. We call this pattern of distortion compression wrinkles. Scale bar, 10 μ m. *b*, Group of dissociated chick heart fibroblasts produce a more complex pattern of substrate distortion, composed of compression wrinkles and tension wrinkles, which radiate from the cell margins and often run between cells. Scale bar, 100 μ m. *c*, Although numerous growth cones extend in a wide band from this explanted dorsal root ganglion, no wrinkles form beneath or adjacent to them. The wrinkles visible are tension wrinkles originating from the two peripheral glial cells. Scale bar, 100 μ m.

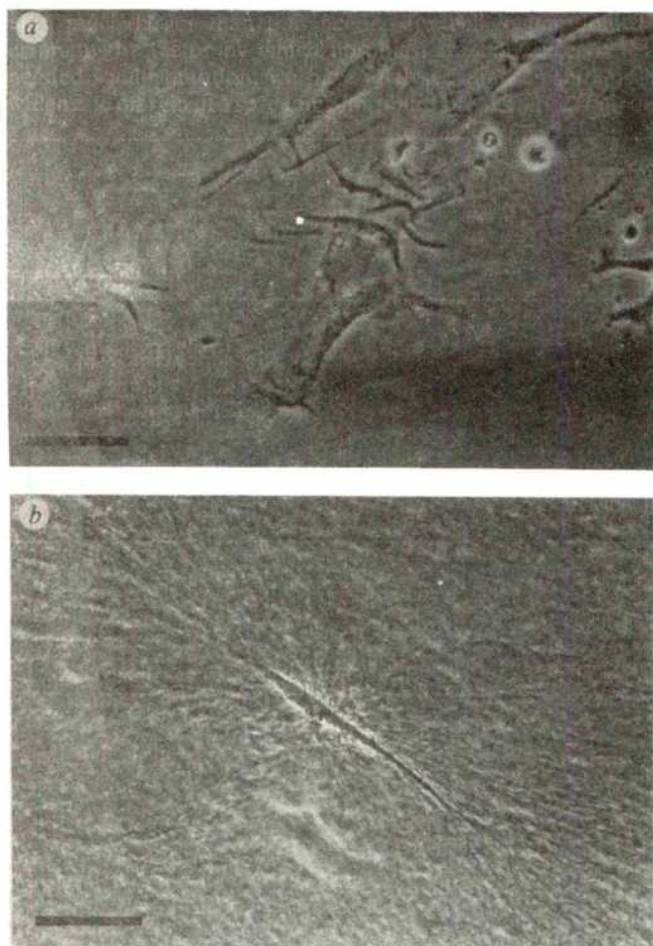


Fig. 2 Collagen substrata are distorted by cellular traction in a pattern similar to the wrinkling of silicone rubber. *a*, Dermal fibroblasts spreading over a thin film of dried collagen produce a pattern of substrate folding, equivalent to compression wrinkling. Collagen solution (Vitrogen 100, Collagen Corporation) was pipetted onto glass coverslips, dried with a stream of sterile air and washed with phosphate-buffered saline (PBS) to remove the acid. Trypsin-dissociated fibroblasts from 7-day chick embryo skin were then seeded onto the collagen films. *b*, Dermal fibroblasts grown in reconstituted collagen gels do not generate wrinkles, but do align the flexible collagen meshwork around themselves. Around the cell body collagen is packed perpendicular to the direction of cell locomotion, while beyond the cell margins, sprays of collagen are aligned parallel to the long axis of the cell. This is equivalent to the compression and stretching seen in Fig. 1*b*. Collagen gels were prepared by mixing collagen solution with 5 × Dulbecco's modified Eagle's medium, fetal calf serum, NaOH and PBS (5:2:1:0.7:1.3 by vol) at 4 °C. Bar, 50 μm.

cells pulled in concert. Weaker still were transformed cells of the lines L-929 and CHO as well as SV40 virus-transformed 3T3 cells. These cells could only deform the rubber visibly where it was already wrinkled or torn. The weakest traction was exerted by macrophages (from mouse peritoneum), polymorphonuclear leukocytes (from human blood) and nerve growth cones (from embryonic chick sensory ganglia). Their weakness was not just a reflection of smaller size. Even when dozens of nerve growth cones lay side by side, their cumulative tractional strength was still too weak to distort our weakest rubber substrata (Fig. 1*c*).

Hence the paradox is that the most mobile cell types exert the weakest traction while the contractile strength of fibroblasts is very much larger than that needed for cell locomotion. To estimate the factor by which fibroblast traction exceeds that needed for locomotion, we have analysed time lapse films in which one margin of a spreading fibroblast breaks loose from the

wrinkled substratum so that it is pulled rapidly sideways by the elastic relaxation of the rubber. Such elastic retractions often displace fibroblasts by distances of many tens or even hundreds of micrometres, distances which correspond to the area of rubber sheet they had compressed beneath themselves while spreading. During these retractions the cell bodies are pulled at speeds of up to 10 μm or more per second, which is much faster than the speed at which the margin of these cells advances during normal locomotion (0.5–1 μm min⁻¹). When we consider that the elastic pull of the rubber cannot be any stronger than the cellular traction which originally distorted it, it becomes clear that the cell's traction is sufficient to propel it 100–1,000 times faster than normal cell locomotion. Thus we conclude that fibroblast traction is at least two or three orders of magnitude stronger than that needed to propel cells of this size at their normal speeds.

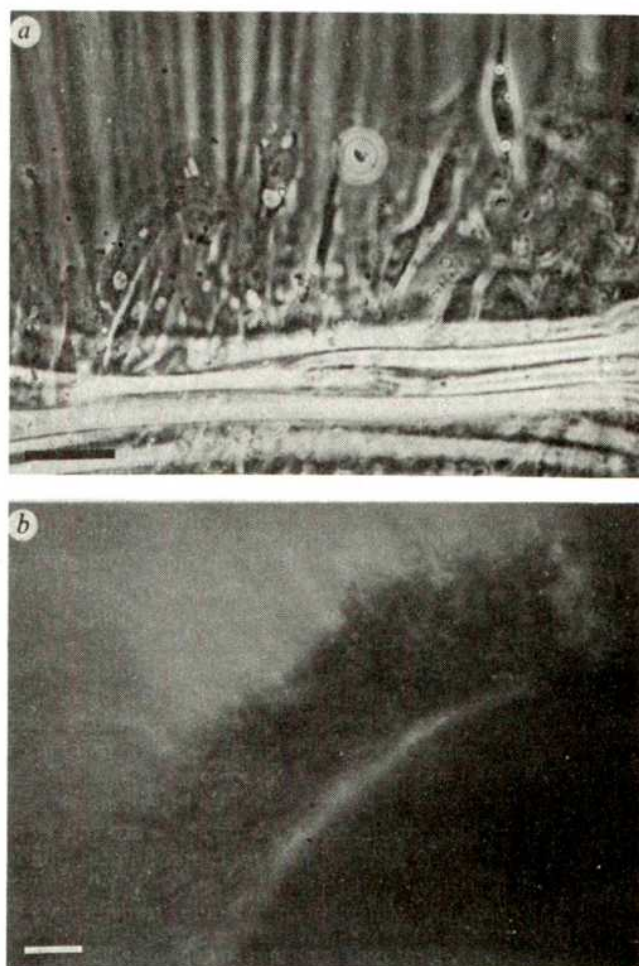


Fig. 3 Comparison of distorting effects of fibroblast traction on silicone rubber sheets and collagen gels. *a*, Marginal outgrowth zone of an explanted 8-day chick ventricle after 48 h in culture on silicone rubber. Adjacent to the explant (bottom of the field) the rubber layer is compressed into a mass of circumferential folds. Beyond the explant the silicone film is stretched resulting in radial wrinkles which continue far beyond the margins of the explant. *b*, The pattern of distortion is equivalent when heart explants are grown in collagen gels and viewed with polarized light. Adjacent to the explant a bright birefringent ring is visible consisting of a capsule of compressed collagen wrapped circumferentially around the explant. Next there is a dark band lacking birefringence where collagen is being compacted and pulled towards the explant by the migrating cells within this area. Collagen fibres farther from the explant are stretched, resulting in a radial pattern of fibre orientation. Scale bars, 100 μm.

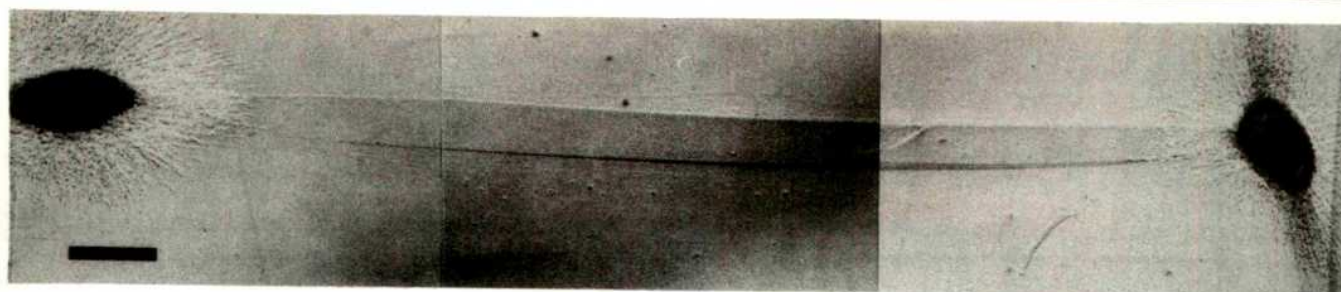


Fig. 4 Two centre effects form in collagen gels even when explants are separated by a distance spanning over 1.5 cm. Collagen fibres become aligned into long axially oriented tracts interconnecting two centres of traction. Heart explants from 8-day chick embryos after 96 h in culture. Scale bar, 1 mm.

Fibroblast traction ought to distort any sufficiently flexible material and we have studied its effects on collagen, the normal substratum of fibroblasts within the body². When aqueous solutions of collagen are air dried to form thin sheets, fibroblasts wrinkle these sheets just as they do rubber substrata (Fig. 2a). When the collagen is reprecipitated as a thick gel, wrinkles are no longer formed. Instead, fibroblast traction compresses and packs the collagen fibres together around and behind the cell bodies. Beyond the cellular extensions which exert this traction, the collagen fibres are stretched into alignment (Fig. 2b). Bell, Ivarsson and Merrill³ have previously described the contraction of collagen gels by fibroblasts. Our observations suggest that these distortions are caused by direct cellular traction—the combined motility and contractility of the cells exerting strong shear forces tangential to their surfaces. Note that this traction is distinct from simple contraction like that of a muscle. For example, the cells elongate instead of shorten as they compress and stretch the collagen around them.

When tissue fragments are explanted into collagen gels, the cumulative traction of many fibroblasts pulls the gel centripetally to form a dense capsule of compressed collagen fibres wrapped around the explant surface (Fig. 3). The region of gel far from the explant is stretched into radial alignment with a 90° rotation in fibre alignment occurring where the fibroblasts exert their traction. Between adjacent explants the pull of this traction aligns the collagen into straight bundles (Fig. 4) which can be more than 1 cm long! The development of these traction fields demonstrates that collagenous structures such as tendons, ligaments, organ capsules and dermal tessellations⁴⁻⁶ need not be laid down by direct secretion of collagen fibres into their eventual arrangement⁷. Instead, fibroblast traction can rearrange and re-pack collagen into these patterns, even beginning from a totally random meshwork.

These traction fields resemble Paul Weiss' centre effects, which lends support to his theories of nerve guidance⁸ but suggests that guidance should be towards foci of strong traction rather than high growth rate. Our findings also point to fibroblast traction as the cause for the apparent contraction of collagen networks around deep wounds, burns and surgically implanted prostheses⁹⁻¹¹.

We had assumed originally that substratum distortion was merely an incidental by-product of cell locomotion, useful for studying its mechanism. However, these results suggest that traction has evolved to serve morphogenetic functions going well beyond cell displacement.

This work was supported by NIH grant GM-24251.

Received 16 October 1980; accepted 21 January 1981.

1. Harris, A. K., Wild, P. & Stopak, D. *Science* **208**, 177-179 (1980).
2. Elsdale, T. & Bard, J. *J. Cell Biol.* **54**, 626-637 (1972).
3. Bell, E., Ivarsson, B. & Merrill, C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1274-1278 (1979).
4. Stuart, E. S. & Moscona, A. A. *Science* **157**, 947-948 (1967).
5. Wessels, N. K. & Evans, J. *Dev. Biol.* **18**, 42-61 (1968).
6. Weiss, P. A. in *Biological Organisation* (ed. Waddington, C. H.) 12-13 (Pergamon, New York, 1959).

7. Trelstad, R. L. & Hayashi, K. *Dev. Biol.* **71**, 228-242 (1979).
8. Weiss, P. A. in *Analysis of Development* (eds Willier, B. H., Weiss, P. A. & Hamburger, V.) 346-401 (Saunders, Philadelphia, 1955).
9. Peacock, E. E. & Van Winkle, W. *Wound Repair* Ch. 3 (Saunders, Philadelphia, 1976).
10. Abercrombie, M., Flint, M. H. & James, D. W. *J. Embryol. exp. Morph.* **4**, 167-174 (1954).
11. Gayou, R. & Rudolph, R. *Ann. Plastic Surg.* **2**, 62-71 (1979).

Genetically restricted cell-mediated cytotoxicity in cattle immune to *Theileria parva*

Elsie M. Eugui & David L. Emery

International Laboratory for Research on Animal Diseases,
PO Box 30709, Nairobi, Kenya

The protozoan parasite *Theileria parva*, which is transmitted by the tick *Rhipicephalus appendiculatus*, produces an acute fatal infection in the lymphoid system of susceptible cattle¹. This disease is a serious constraint to livestock improvement and production in large areas of East Africa¹. The parasite invades host lymphocytes, inducing rapid proliferation followed by widespread lymphocytolysis². Cattle which recover from theileriosis (East Coast fever) spontaneously, or which are immunized by infection and treatment with tetracycline³, are resistant to reinfection with the same isolate of *T. parva* for at least 3 yr⁴. Immunity against infection with the parasite cannot be ascribed to the production of specific antibodies⁵, but can be transferred adoptively between twins with thoracic duct leukocytes from the immunized partner⁶. These observations suggest that protective immunity is associated with cell-mediated mechanisms. We have now examined the capacity of leukocytes from immune cattle to lyse parasitized lymphoblastoid and non-parasitized tumour cell lines either directly or after stimulation in an autologous mixed lymphocyte reaction (MLR). In contrast to the nonspecific lytic activity of leukocytes from immune cattle reported by Pearson *et al.*⁷, we describe the sequential appearance in the lymph and blood of immune cattle, of cytotoxic leukocytes with activity restricted to target cells carrying the autologous genotype. These observations suggest that a major component of protective immunity to *T. parva* is mediated by cytotoxic cells which lyse parasitized cells in a genetically restricted fashion. We have also found that during a primary infection with *T. parva* cytotoxicity was manifested against allogeneic parasitized cells and xenogeneic uninfected target cells, but not against autologous infected cells. The features of cell-mediated immunity to *T. parva* during primary infection and immunization are discussed more fully elsewhere⁸.

Studies of cell-mediated immunity in cattle and other outbred species have been hampered by the lack of syngeneic cell lines for monitoring the development of cytotoxic responses. In the case of *T. parva*, this limitation has been overcome by using the

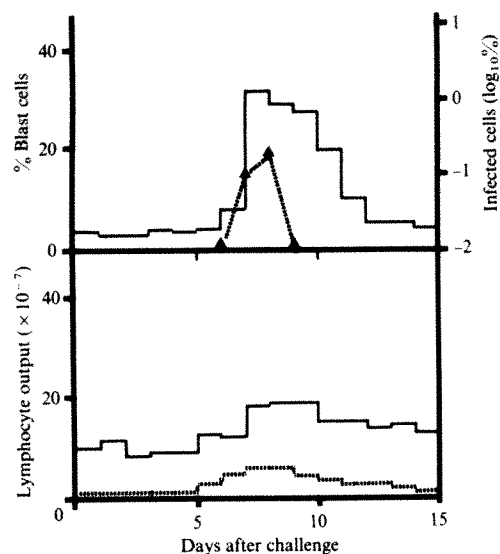


Fig. 1 The efflux in central lymph of normal and infected lymphocytes following challenge of three immune calves with fresh sporozoites adsorbed on autologous PBL. Results are expressed as mean average total (—) and blast lymphocyte output per hour (---); mean average percentage blast cells (—▲—▲) and mean average percentage of lymphocytes infected with macrophages of *T. parva* (—▲—▲), as determined by direct immunofluorescence¹¹.

ability of sporozoites of the parasite to transform *in vitro* leukocytes from normal cattle into lymphoblastoid cell lines infected with macrophages of *T. parva*⁹. We examined the specificity of cytotoxic responses of cattle immunized with *T. parva* using autologous and allogeneic infected target cells derived from such *in vitro* transformations, as well as xenogeneic non-parasitized cells (YAC-1). The latter is a mouse tumour cell line susceptible to lysis by natural killer (NK) cells of murine origin¹⁰, and we have found that it is lysed by normal bovine peripheral blood leukocytes (PBL).

Freisian calves aged 6–8 months, serologically negative for *T. parva*, were immunized by subcutaneous inoculation with sporozoites of *T. parva* (Muguga) and a concurrent intramuscular inoculation with long-acting tetracycline (20 mg per kg., Pfizer)³. Such immune animals were challenged subcutaneously 4 weeks later with the same isolate of the parasite from stablate or as fresh sporozoites adsorbed on autologous PBL. At intervals after challenge, cytotoxic responses were monitored in PBL. After a further 6 weeks, three calves received a subcutaneous inoculum of 5×10^8 autologous infected cells, and two calves were inoculated with another dose of infective sporozoites. Again, cytotoxicity was evaluated in PBL at intervals after the second challenge.

Table 1 Specificity of the cytotoxic response in immune animals after challenge with 5×10^8 autologous *T. parva*-infected cells

PBL donor	Target cell, <i>Theileria</i> -infected	% Specific lysis* (days after challenge)	
255	Th-256	3	7
	Th-588	2.1	48.2
	Th-589	1.0	0
	Th-589	0	0
588†	Th-588	5.3	70.1
	Th-589	0	39.0
	Th-256	3.8	2.2
	Th-589	7.1	39.7
589†	Th-588	4.5	51.5
	Th-256	4.7	0
	Th-589	5.0	5.3
Normal‡	Th-588	5.2	4.3
	Th-589	2.3	0.5

* Ratio of effector to target cells was 100:1 in macroassays (Fig. 2 legend).

† Animals 588 and 589 are chimaeric twins.

‡ Results are average of two normal control animals tested the same day as the immune challenged animals.

When immune animals were challenged with sporozoites as stablate, macrophage-bearing cells were not detected in biopsies of the draining lymph nodes during the ensuing 3 weeks, nor in lymph efferent from these nodes. Lymph node cells (LNC) from the draining lymph nodes were unable to stimulate an autologous MLR using normal PBL cryopreserved before the primary infection, whereas they stimulated allogeneic

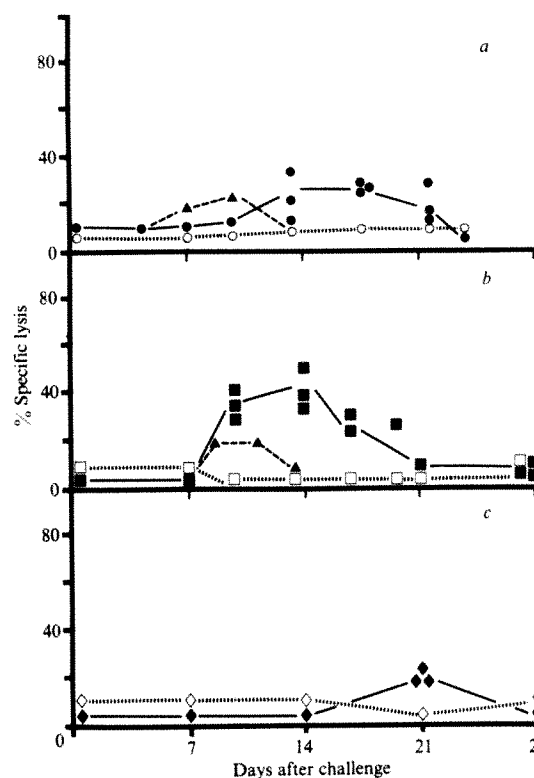


Fig. 2 Three calves were immunized as described and were inoculated subcutaneously 4 weeks later with a lethal dose of sporozoites of *T. parva* (Muguga) from stablate. At intervals after challenge, PBL were purified using Ficoll-paque¹¹ (Pharmacia) and were tested in standard 4-h microcytotoxicity⁷ and macrocytotoxicity assays¹⁶ against autologous and allogeneic parasitized cell lines as well as a xenogeneic tumour cell line (YAC-1). All cells were suspended in RPMI-1640 culture medium (Flow Labs) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS), 20 $\mu\text{g ml}^{-1}$ glutamine and 10 mM HEPES buffer. For the microcytotoxicity assay, 5×10^4 target cells (in 50 μl) which had been previously labelled with 200 μCi ⁵¹Cr (Radiochemical Centre) for 1 h at 37 °C, were added in duplicate to 200 μl of effector cells in flat-bottomed microtitre plates (Linbro 96-TC), at effector : target (E : T) ratios of 80 : 1, 40 : 1, 20 : 1 and 10 : 1. Plates were incubated for 4 h at 37 °C in a humidified atmosphere of 5% carbon dioxide in air. Macroassays were performed in triplicate in plastic tubes, where 4×10^4 target cells (labelled with 150 μCi ⁵¹Cr for 30 min at 37 °C) were added to effector cells at E : T ratios of 100 : 1 and 50 : 1 in a total volume of 1 ml. The tubes were incubated as above for 12 h. Appropriate controls were included in both assays to determine total and spontaneous release of isotope. After incubation, 100 μl of supernatant was carefully collected from the microassay and 500 μl from the macroassay and counted in a γ -scintillation spectrometer (Packard). Results are expressed as the percentage of specific lysis calculated from the formula:

$$\% \text{ specific lysis} = 100 \times \frac{\text{c.p.m. in test} - \text{c.p.m. in spontaneous release}}{\text{total release c.p.m.} - \text{c.p.m. in spontaneous release}}$$

In addition, unidirectional MLRs were established using responder PBL from immune calves and stimulator cells from autologous infected cell lines which had been treated with 25 $\mu\text{g ml}^{-1}$ mitomycin C (Sigma) for 30 min at 37 °C. 4×10^7 PBL were added to 8×10^6 stimulator cells in 16 ml of culture medium (as above) supplemented with 10^{-4} M 2-mercaptoethanol, 100 IU ml^{-1} penicillin and 100 $\mu\text{g ml}^{-1}$ streptomycin. Aliquots (2 ml) were dispensed into 24 well $\times$ 16 mm culture plates (Linbro) and incubated at 37 °C in a humidified atmosphere of 5% CO₂ in air. After 5 d, the viable cells were recovered, washed and used in the microcytotoxicity assay. The figure depicts results from macrocytotoxicity test (a●) (50:1) direct (b■) and MLR-generated (c●) microcytotoxicity assays (40:1) using PBL effector cells against autologous (—) and allogeneic infected target cells (—○—), and efferent lymphatic lymphocytes (ELL) as effector cells against autologous infected targets (—▲—▲).

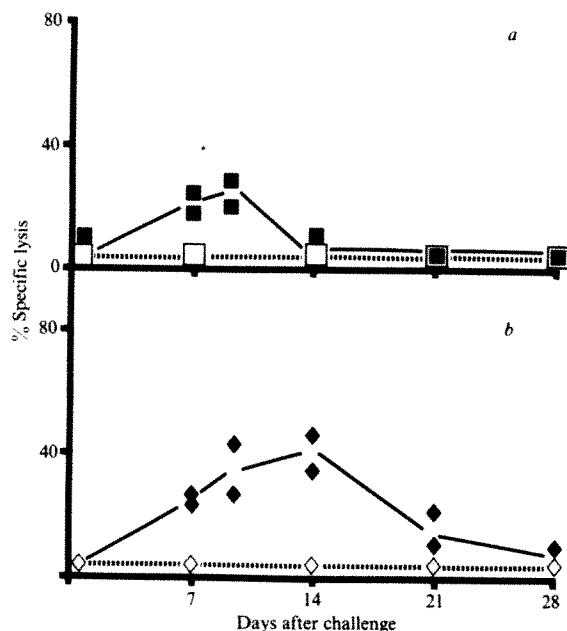


Fig. 3 Cytotoxic response of PBL from two immune calves after a second challenge with stablilated infective sporozoites of *T. parva* which were inoculated subcutaneously 10 weeks after infection and treatment and 6 weeks after the first sporozoite challenge. Results are depicted direct (a) and MLR-generated (b) microassays (40:1) using as targets autologous (—) and allogeneic (---) infected cell lines.

MLRs (data not shown). This result contrasted with the substantial autologous MLRs initiated by LNC during the second week of a lethal primary infection¹¹. When the challenge consisted of freshly isolated sporozoites adsorbed on autologous PBL, a transient parasitosis (<0.5%) was observed in efferent lymph from the draining lymph node on day 9. In these calves, the proportion of blast cells increased sharply from 5 to 30% on days 8–9 after challenge (Fig. 1). Cells from lymph nodes in which parasites were detected induced MLRs in normal cryopreserved autologous PBL. These results indicated that LNC are able to stimulate MLRs only when blasts, presumably infected cells, are present. Moreover, in immune animals challenged with sporozoites in the form of stablilate, the infection did not progress to a stage where macroschizonts could be detected by immunofluorescence or where infected cells could stimulate an autologous MLR.

Both challenge protocols with infective sporozoites elicited the development of cytotoxicity against autologous infected cells (Fig. 2a, b). Lytic activity was first detected in PBL 10–12 days after challenge, was maximal around day 14 when 20–50% of specific lysis was observed, and declined thereafter. Unlike in primary infections, no cytotoxicity was demonstrable against allogeneic cells or xenogeneic target cells at any stage. In one animal, lymphocytes draining from the regional lymph node exhibited cytotoxic activity against autologous infected cells from days 9–14, after the parasitosis had disappeared and before the detection of lytic activity in PBL (Fig. 2a, b). Calves re-challenged with a second inoculum of sporozoites as stablilate also showed cytotoxic responses in PBL which were again restricted to autologous target cells (Fig. 3a). These appeared at day 7, earlier than after the initial sporozoite challenge.

Similarly, strong cytotoxicity against autologous infected cells was found in three immune calves 6–14 days after challenge with 5×10^8 autologous infected cells (Fig. 4a, b). This lytic activity, which ranged between 20 and 60%, declined during the third week after challenge; once again, no activity was detected against allogeneic infected cells or xenogeneic tumour target cells. Moreover, the cytotoxicity was not due to any unique sensitivity to lysis of the allogeneic and autologous cell lines used: in experiments using crossed combinations of effector and target cells simultaneously, cytotoxicity was consistently observed against autologous but not allogeneic infected target cells

(Table 1). After challenge with autologous parasitized cells, leukocytes from two immune twin calves killed infected cells of their respective partners as efficiently as autologous cells (Table 1).

In other experiments, MLRs were performed using PBL from immune animals as responders and mitomycin-treated autologous infected cells as stimulators. When PBL were taken 21 days after challenge of immune animals with sporozoites and cultured for 5 days, a relatively low peak of cytotoxicity (~20%) was observed (Fig. 2c). When PBL were taken from immune cattle 7–24 days after challenge with autologous infected cells or with a second dose of stablilate, and cultured for 5 days, stronger cytotoxicity was detected (30–80%) (Figs 3b, 4c). In both cases, lysis was confined to autologous infected targets.

The present observations provide evidence that a population of specific cytotoxic cells is developed in cattle which have recovered from a primary infection with *T. parva*. These cells have the capacity to lyse autologous or cotwin but not allogeneic parasitized cells. The absence from peripheral blood of cytotoxic cells, and of precursors from which cytotoxic cells can be generated in MLR, at intervals other than those defined in this study, probably indicates that such cells reside in the solid lymphoid organs between encounters with the parasite. Following successive challenge with *T. parva*, specific cytotoxic cells enter central lymph and blood with a latent period progressively declining to a minimum of 6–7 days. Our findings demonstrate that cytotoxicity is effected directly by leukocytes collected from blood and lymph, discounting the necessity for an initial MLR to detect lytic activity⁷.

The genetic restriction of the cytotoxic response against *T. parva* in immune animals reported here contrasts with the lysis of allogeneic target cells reported by Pearson *et al.*⁷. The reasons for the discrepancy are not clear, but may reside in the multiple inoculations of autologous infected cells that those investigators

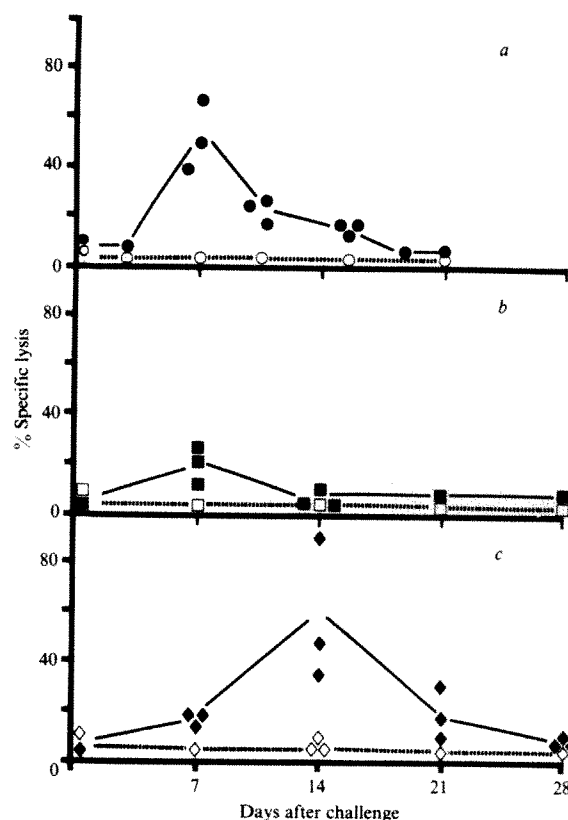


Fig. 4 Cytotoxic response of PBL from three immune calves at intervals after challenge with 5×10^8 autologous infected cells which were inoculated subcutaneously 10 weeks after immunization and 6 weeks after challenge with stablilated *T. parva*. Results are depicted from direct macrocytotoxicity tests (a) (50:1), direct (b) and MLR-generated (c) microassays (40:1) using as targets autologous (—) and allogeneic infected cell lines (---).

used to obtain PBL which expressed cytotoxicity after an MLR. Repeated challenges of immune calves with parasitized autologous cells in our studies has only yielded cytotoxicity of autologous infected target cells. After challenge of each of 18 immune animals so far examined, killing of allogeneic targets has not been detected in direct assays against a range of parasitized allogeneic cell lines. Alternatively, the former workers used only a single immune animal and one allogeneic infected cell target, and it was not established that this particular combination was genetically unrelated; effector and target could have shared some of the relevant histocompatibility antigens.

The restriction of cytotoxicity by the cells of immune animals to autologous or chimaeric twin partner infected target cells is analogous to the restriction of T-lymphocyte-mediated cytotoxicity of murine virus-infected cells¹², or hapten-modified syngeneic targets¹³. This is the first time that such restriction has been observed for a protozoan parasite. By analogy, cytotoxic lymphocytes from immune cattle may recognize complexes of parasite-specific antigens and polymorphic host determinants (probably major histocompatibility antigens), on the surface of infected cells. Analysis of this interpretation will require more information about products of the major histocompatibility complex in cattle than is yet available. The kinetics of appearance and activity of specific cytotoxic cells suggest that these cells represent a major mechanism of recovery from infection. They are observed in animals that are being immunized by infection and treatment, and in immune animals challenged with sporozoites or macroschizont-infected cells. Recovered animals that are immune to the homologous isolate of the parasite are susceptible to some other isolates of *T. parva*¹⁴. If specific cytotoxicity is correlated with resistance, the effector cells would be expected to kill autologous cells infected with isolates of the parasite to which there is cross protection *in vivo*. Experiments are in progress to clarify this point, with the objective of developing a laboratory test to ascertain whether isolates or antigenic types of *T. parva* might be cross-protective *in vivo*, an important prerequisite for a vaccination programme.

One approach to immunoprophylaxis against *T. parva* envisages using infected cell lines. With respect to this approach, genetic restriction of the cellular response to *T. parva* implies that successful prophylaxis requires the parasite to become established in the lymphocytes of the host. The factors which govern the transfer of *T. parva* macroschizonts from allogeneic to host cells have not been elucidated. Although this exchange can occur, the process seems to be relatively inefficient, which may account for the large number (of the order of 10⁸) of allogeneic cells required to infect or immunize cattle. In contrast, infections can be initiated with as few as 100 autologous macroschizont-bearing cells¹⁵. Studies are in progress in this laboratory to characterize the antigenic heterogeneity of the various isolates of *T. parva* and to define the nature of the genetic restriction involved in the acceptability of parasitized allogeneic cells by the host for vaccination. Cellular immunity to parasites has previously been attributed to production by lymphocytes of factors that activate macrophages so that they limit the multiplication of parasites within them. We now describe a new mechanism: direct killing by lymphocytes of parasitized cells. This may well also apply in other intracellular parasite infections, along with the genetic restriction defined here.

We thank Dr G. Büscher for theilerial sporozoites, Messrs R. Nelson for *Theileria*-infected cell lines and Tom Tenywa for technical assistance, Pfizer Laboratories for the provision of Terramycin LA, Drs A. C. Allison, S. Black, W. I. Morrison and Max Murray for constructive discussions and support throughout these investigations, and Ms Leah Agina for typing the manuscript.

Received 28 August; accepted 17 December 1980.

1. Barnett, S. F. *Bull. Epizoot. Dis. Afr.* **5**, 543-557 (1957).
2. De Martini, J. C. & Moulton, J. F. *J. comp. Path.* **83**, 281-298 (1973).
3. Radley, D. E. *et al. Vet. Parasit.* **1**, 51-60 (1975).

4. Burridge, M. J., Morzaria, S. P., Cunningham, M. P. & Brown, C. G. D. *Parasitology* **64**, 511-515 (1972).
5. Wagner, G. G., Duffus, D. P. H. & Burridge, M. J. *Parasitology* **69**, 43-53 (1974).
6. Emery, D. L. *Res. vet. Sci.* (in the press).
7. Pearson, T. W., Dolan, T. T., Stagg, D. S. & Lundin, L. B. *Nature* **281**, 678-680 (1979).
8. Emery, D. L., Eugui, E. M., Nelson, R. & Tenywa, T. *Immunology* (in the press).
9. Brown, C. G. D., Stagg, D. A., Purnell, R. E., Kanhai, G. K. & Payne, R. C. *Nature* **245**, 101-103 (1973).
10. Cikes, M., Friberg, S. Jr & Klein, G. J. *Natn. Cancer Inst.* **50**, 347-350 (1973).
11. Emery, D. L. & Morrison, W. I. *Immunology* **40**, 229-237 (1980).
12. Doherty, P. C. & Zinkernagel, R. M. *J. exp. Med.* **141**, 502-507 (1975).
13. Shearer, G. M., Rehn, G. R. & Garbarino, C. A. *J. exp. Med.* **141**, 1348-1356 (1975).
14. Purnell, R. E. *Adv. Parasit.* **15**, 83-132 (1977).
15. Büscher, G., Morrison, W. I. & Nelson, R. (in preparation).
16. Eugui, E. M. & Allison, A. C. *Bull. Wild Hlth Org.* **57**, Suppl. 1, 231-238 (1979).

Normal levels of natural cytotoxic cells against solid tumours in NK-deficient beige mice

Osias Stutman & Michael J. Cuttito

Cellular Immunobiology Section, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

Natural cell-mediated cytotoxicity (NMC) capable of *in vitro* lysis of various lymphoid and non-lymphoid tumours has been described in mice and other species, including man¹. NMC has been proposed as a first level of defence against tumour growth *in vivo*, one which does not need the priming of the conventional immunological response¹. The effector cells of NMC seem to belong to a special category of lymphoid cells, being neither classical T or B cells nor macrophages¹⁻⁵; natural killer (NK) cells have been proposed as the prototype effector cell¹, although some heterogeneity among effector cells seems to exist, depending on the target cells used for testing⁴⁻¹⁰. Two main subgroups of NMC effector cells have been defined: NK cells directed against lymphoma targets^{2,3,6,7} and natural cytotoxic (NC) cells directed against solid non-lymphoid tumours^{4,5,8-10}. We describe here another distinction between the two systems: while NK activity is low in mice homozygous for the beige (*bg*) gene¹¹⁻¹³ NC activity in spleen cell preparations from these animals is comparable with that observed in the appropriate controls (*bg*/+ and +/+ littermates). The *bg* syndrome of mice affects lysosome, melanosome and enzymatic functions and is a homologue of the Chediak-Higashi syndrome of man¹⁴. Defective NK activity in blood lymphocytes has also been reported in patients with Chediak-Higashi syndrome¹⁵. We also show that several mouse strains which have low NK activity, have normal or high levels of NC functions, expanding our previous observation that NC and NK cells are under distinct genetic control^{4,16}.

Table 1 summarizes the similarities and differences between NK and NC effector cells. As postulated earlier, it seems that NK and NC cells belong to a related, though distinct, family of effector cells, detectable in normal mice, which are capable of *in vitro* lysis of tumour cells without prior stimulation^{4,5,8}.

Table 2 shows our measurements of NK and NC activity of spleen cells against YAC-1 lymphoma and Meth-A sarcoma targets respectively, using C57BL/6J *bg*¹/*bg*¹ (line 4) and various controls (lines 1-3). Clearly NC activity is comparable in all groups studied, including the homozygous *bg* animals, while NK activity was detectable but significantly lower in the homozygous *bg* mice, in accordance with published reports¹¹⁻¹³. A small group of SB/Le mice which carry the *bg* gene derived from a radiation-induced mutation (the *bg*¹ being a spontaneous mutation)^{17,18} was also studied (lines 5 and 6). While NC levels between the homozygous *bg* and controls were comparable, NK activity was again markedly reduced in the homozygous *bg* (line 6). F₁ hybrids between C57BL/6J *bg*¹/*bg*¹ and SB/Le *sa bg*/*sa bg* also showed the same low NK reactivity (line 8), supporting the view that both *bg* mutations are acting on the same gene^{17,18}, and that the NK defect is not associated with the satin (*sa*) gene¹⁷.

In addition, Table 2 (lines 9–13) shows that other mouse strains such as A/J and some sublines³, SJL/J (S. Kaminsky and G. Cudkovic, personal communication) and I/St mice, which show low levels of NK activity when tested against YAC-1 targets, have normal levels of NC function against Meth-A sarcoma targets ('normal' means comparable with the syngeneic BALB/c strain⁴). Furthermore, I/St mice belong to the group of mouse strains with the highest NC reactivity against Meth-A targets^{4,8}.

The observation that transplanted tumours grow better and produce more metastases in homozygous *bg* mice^{19–21} supports the view of a possible *in vivo* surveillance role^{2,3} of the NK cells. On the other hand, one study, using a small group of homozygous *bg* and controls (8 mice per group), also showed that the incidence and time of appearance of sarcomas after administration of 3-methylcholanthrene were comparable²¹. This suggests that such 'normal' tumour incidence after chemical carcinogenesis in *bg* mice may be due to the adequate levels of NC cells present in such animals, as described here. Much larger groups of animals, as well as different carcinogen dosages have to be tested in homozygous *bg* mice and controls before this issue is settled. A similar speculation was made concerning the 'normal' incidence of spontaneous and induced tumours in athymic nude mice²², which have high levels of both NK^{2,3} and NC^{4,16} activity. The single published study on incidence of spontaneous tumours in *bg* mice¹⁷ showed that while the homozygous SB/Le *bg* animals had increased susceptibility to pneumonia, the incidence of lymphomas or other tumours was not increased. In fact, the incidences of lymphoproliferative lesions and lymphomas seemed to correlate more with the recessive gene *sa* than with *bg*¹⁷. Some cases of Chediak-Higashi syndrome in man develop a lymphoma-like tumour²³.

Similar comments can be made concerning the other mouse strains with low NK but normal NC activity described in Table 2 (lines 9–13). With the exception of SJL mice, which develop a high incidence of a reticulum cell lymphoid neoplasm²⁴, none of the other strains can be catalogued as having high tumour incidence^{24–27}. A/J mice show high incidence of lung adenomas at 18 months of age^{24,25} and mammary tumours in breeding females^{24,25} which are comparable with other mouse strains with normal NK activity. I/St mice in addition, are quite resistant to polycyclic hydrocarbon carcinogenesis^{26,27}. Again, as speculated, the 'normal' incidence of the common mouse tumours²⁴ in these strains with low NK activity (and also in the case of A/J mice with a deficient ability to generate tumoricidal macrophages²⁸) may be due to the normal NC activity observed in these animals. The low NK activity in the SJL may contribute to the high incidence of lymphoid tumours in such animals. On the other hand, additional factors may also be operative in SJL mice, as these animals were specifically bred for high incidence of the lymphoid tumours²⁴, and the incidence of spontaneous lymphomas is not clearly increased in mouse strains with low NK activity such as A sublines and 129/J (see ref. 3 for NK information and refs 24 and 27 for lymphoma incidence). Anti-cancer surveillance theories (whatever the mechanism for such surveillance) have predicted that an increase in tumour incidence would be the consequence of any spontaneous or induced interference with the surveillance device²⁹, but this was never clearly fulfilled for any surveillance model in animals or man^{27,30}.

It is tempting to speculate that NK–NC cells represent some form of alternative pathway of immunological surveillance ('alternative' meaning that it is not mediated by T cells as originally proposed in the theory²⁹, but that it is serving the same purpose of handling incipient tumours). Thus, animal models showing selective deficits of NK activity may become important for the determination of the *in vivo* role of such cells. However, using the same type of argument that we used in our criticism of the T-cell dependency of immunological surveillance^{22,27,30}, it may be that the effects detected in the NK–NC systems against tumours are only by-products of effector systems which serve other functions *in vivo*, perhaps related either to regulation of

Table 1 Comparison between NK and NC effector cells of natural cell-mediated cytotoxicity

Similarities	Differences
Pre-existing high levels in normal hosts, no need for apparent priming	Mouse strain distribution of activity
No evidence of immunological memory	Age of appearance
Levels of activity influenced by interferon	Affect of <i>in vivo</i> treatment with ⁸⁹ Sr, oestrogen, cyclophosphamide
Type of activity (that is, activity against otherwise negative targets) influenced by interferon and interferon inducers	Surface determinants (FcR, Thy-1, NK 1, Qa 2 and 5, Asialo GM1)
No H-2 restriction for cytotoxicity	
High reactivity dominant in F ₁ hybrids	
Present in athymic nude mice	
Tissue distribution	
No surface Ig and Ia negative	
Non-adherent	
Non-phagocytic	
Size	
<i>In vivo</i> effects of irradiation and steroids	
Blocking of cytotoxicity by monosaccharides	

NC^{4,5,8–10} and NK³ cells were compared using mostly NK data with YAC-1 targets. Some of the properties suggesting differences are not general: NK cells against EL4 lymphoma targets are not affected by *in vivo* treatment with ⁸⁹Sr (ref. 6), as are NC cells⁸; age of appearance of NC activity is usually earlier than NK, although that is dependent on the target cells used and in some cases activity is detected only at 3–4 weeks of age, as with NK⁹; with surface determinants, although NC cells are negative for those listed above^{5,8}, note that while some markers like Asialo GM1 seem to be present on all NK effector cells³⁴, other markers such as Thy-1, FcR or Qa 5 may be present only on fractions or subpopulations of NK cells^{37,38}. The surface phenotype of most NK cells is NK 1 (refs 7, 38), Qa 2 (refs 38), Qa 5 (ref. 38) and Asialo GM1 (ref. 34) positive, while fractions of NK cells also express Thy-1 (refs 37, 38) and receptors for Fc (refs 2, 3). All these markers are absent from NC cells^{5,8,39}. Other markers present on NK cells not yet tested in NC include Ly 5, Ly 6 and receptors for *Helix pomatia*^{1–3,38}. Some of the similarities also deserve comment. For example, NC cells show a certain degree of adherence to nylon wool⁵ which was also observed with certain types of activated NK cells^{1–3}, although the bulk of NK activity is non-adherent to nylon. The blocking of NCMC by monosaccharides¹⁰ also shows differences: while NC cells show some selectivity to blocking by D-mannose (and galactose), NK cells show less selectivity and other physiological sugars can also block NCMC. For a review see ref. 39.

haematopoiesis, defence against virus or other infectious or parasitic agents^{1–3,8,14}.

Note three points concerning the present comparisons: (1) The NK defect observed in beige mice, although profound, is not total, and some degree of reactivity can be observed either in long-term assays (G. Cudkovic, personal communication) or after *in vivo* administration of interferon inducers^{19,31,32}. Thus, the present comparisons using a 4-h assay for NK but a 24-h assay for NC cells can be questioned. Kinetic studies of NC-mediated cytotoxicity have shown that most of the activity occurs within the first 12 h (refs 5, 8), there is no 'activation' of NC (or NK) cytotoxicity after 24 h in the assay and there is undetectable production of interferon during the 24-h assay (E. Lattime, W. Stewart III and O. S., unpublished results), suggesting that NC cells are not a special kind of *in vitro* activated NK cell. In addition, the differences in the expression of surface antigens between NC and NK cells also argues against such an interpretation, especially as activated NK cells also express such markers as Qa 5 and Asialo GM1 (refs 33, 34). (2) Although the beige mice seem to be defective solely in NK-mediated cytotoxicity^{11–13}, *in vitro* antitumour cytostasis and cytotoxicity by *in*

Table 2 NC and NK activity of spleen cells from C57BL/6J, homozygous *bg*^J, heterozygous *bg*^J and controls, and from other mouse strains with low NK activity (A/J and sublines, SJL/J and I/St)

Spleen cell donors	No. of mice tested	NC activity anti-Meth A			Nk activity anti-YAC-1		
		100	50	10	100	50	10
1. C57B/6J +/+	12 (12)	42 ± 5.1	33 ± 5.9	16 ± 6.3	37 ± 2.5	21 ± 1.3	14 ± 1.0
2. C57BL/6J <i>bg</i> ^J /+	3 (3)	42 (38–48)*	37 (33–41)	15 (15–16)	34 (31–38)	25 (20–33)	19 (12–28)
3. C57BL/6J <i>bg</i> ^J /+ and +/+ littermates	16 (12)	44 ± 4.8	34 ± 5.7	22 ± 5.5	30 ± 1.7	23 ± 1.5	13 ± 1.1
4. C57BL/6J <i>bg</i> ^J / <i>bg</i> ^J	16 (16)	41 ± 3.9	32 ± 4.4	20 ± 4.6	14 ± 1.1†	8 ± 0.9	4 ± 0.6
5. SB/Le +/+ and <i>bg</i> <i>sa</i> /++	5 (5)	53 ± 4.9	40 ± 4.7	27 ± 3.0	29 ± 1.9	19 ± 1.8	9 ± 1.1
6. SB/Le <i>bg</i> <i>sa</i> / <i>bg</i> <i>sa</i>	5 (5)	57 ± 4.7	33 ± 3.5	25 ± 2.8	4 ± 0.6	2 ± 0.6	0
7. (C57BL/6 × SB/Le)F ₁ +/+	10 (5)	55 ± 5.1	40 ± 4.1	29 ± 3.0	38 ± 2.3	25 ± 1.9	17 ± 1.3
8. (C57BL/6J × SB/Le)F ₁ <i>bg</i> ^J / <i>bg</i> <i>sa</i>	10 (9)	54 ± 4.3	42 ± 3.9	27 ± 2.9	7 ± 1.0	4 ± 0.6	2 ± 0.1
9. A/J	11 (6)	49 ± 2.5	26 ± 3.4	15 ± 1.4	11 ± 1.7	6 ± 1.1	5 ± 1.1
10. A/Boy	10 (5)	42 ± 3.3	26 ± 3.4	10 ± 1.2	10 ± 1.8	5 ± 0.7	2 ± 0.8
11. A.T1a ^b /Boy	10 (5)	39 ± 2.6	20 ± 4.3	17 ± 1.6	12 ± 3.4	9 ± 2.7	4 ± 1.5
12. SJL/J	6 (6)	39 ± 4.0	27 ± 2.6	10 ± 1.1	5 ± 0.9	3 ± 1.5	0
13. I/St	17 (10)	63 ± 3.0	39 ± 4.1	24 ± 2.7	5 ± 0.8	2 ± 0.7	0

All mice were 6–8 weeks old when used as spleen cell donors, and in most cases were simultaneously tested for NC and NK activity. The beige mice, controls and A/J mice were obtained from Jackson Laboratories; the SJL/J mice were a gift from Dr G. Cudkovic; the A and its congenic pair A/T1a^b were obtained from Dr E. A. Boyse; the I/St mice were from our colonies⁴, as were the *bg* F₁ hybrids and controls. All values presented are mean ± S.E.M. of individual mice. Numbers in parentheses indicate number of mice tested simultaneously for NK activity. For details on the BALB/c sarcoma Meth A and the A/Sn lymphoma YAC-1 see refs 4, 5, 8, 10. NC activity was tested using ³H-proline-labelled adherent target cells, 100, 50 and 10 effector/target ratios and 24-h incubation as described in refs 4, 5, 8. NK activity was tested with ⁵¹Cr-labelled YAC cells in suspension, using the same effector/target ratios in 4-h release assays as described in ref. 10.

* As only three animals were tested, the results are presented as mean and range in parentheses.

† One animal had 32, 21 and 20% cytotoxicity at the three effector/target ratios tested. All the others had values of 15% or less at the 100:1 ratios.

in vivo activated macrophages is defective³⁵, while activity of *in vitro* activated macrophages and promonocytes of *bg* mice may be normal^{11–13,35}. Thus, the beige mice may have additional defects of cells, such as macrophages, putatively involved in defence against tumours, making *bg* an interesting model of NCMC deficiency, but certainly not the litmus test for definition of NK function *in vivo*. (3) The apparent tumour 'selectivity' of the NC–NK system for solid tumours versus lymphomas deserves comment. One proposed terminology even uses NK_L for the conventional antilymphoma activity and NK_S or NK_{solid} for a cell with similar characteristics as NC, although tested in a different assay⁷. 'Solid' is used throughout, in the traditional pathologist's sense, opposing it to the diffuse malignancies such as leukaemia, and not defining a particular state of matter. Such tumour-type preference of NC and NK may not be absolute. Metastatic spread to the lungs of any type of tumour may be controlled by NK-like cells^{19,36}, while tumour growth in other sites may be controlled by NC-like mechanisms, at least for certain solid tumours. The *in vivo* studies with beige mice show a deficient response to lymphomas²⁰ and melanomas¹⁹ (technically a solid tumour, although 'modified' to be sensitive to NK cytotoxicity⁷ by long-term *in vitro* passage¹⁹). Such deficiency is evidenced by an increase in local growth of the lymphomas²⁰ and an increase in spontaneous and artificial lung metastasis from solid tumours^{19,21}. On the other hand, local tumour growth of melanoma¹⁹ or the Lewis lung carcinoma³⁵, as well as the appearance of lymph node melanoma metastases¹⁹ were comparable in beige mice and controls. It is also tempting to speculate that the differences between tumour type and NC or NK cytotoxicity are due to the predominance on certain tumours (lymphomas compared with solid tumours) of determinants that are recognized by NC and NK cells respectively, perhaps related to either concentration or appropriate display of certain monosaccharides associated with membrane glycoproteins or glycolipids, as may be suggested by our studies of blocking of cytotoxicity, showing D-mannose preference for the NC system¹⁰.

Thus the beige mutation which affects NK cells has no effect on NC activity, a situation observed in other low-NK mouse strains such as A/J and sublines, SJL/J and I/St. Whether this situation explains the overall low tumour incidence of most of these strains, if indeed NCMC is considered as an antitumour surveillance device, needs further study.

This research was supported by grants CA-08748, CA-15988 and CA-17818 from the NIH and grant IM-188 from the American Cancer Society. We thank Mary L. Devitt for the animal care, Linda Stevenson for preparation of the manuscript, Drs Edmund Lattime, Stuart Macphail and Richard Miller for criticism and Dr G. Cudkovic for providing unpublished results.

Received 21 October 1980; accepted 19 January 1981.

1. Heberman, R. B. (ed.) *Natural Cell-mediated Immunity Against Tumors* (Academic, New York, 1980).
2. Herberman, R. B. & Holden, H. T. *Adv. Cancer Res.* **27**, 305–377 (1978).
3. Kiessling, R. & Wigzell, H. *Immun. Rev.* **44**, 165–208 (1979).
4. Stutman, O., Paige, C. J. & Feo Figarella, E. *J. Immun.* **121**, 1819–1826 (1978).
5. Paige, C. J., Feo Figarella, E., Cuttito, M. J., Cahan, A. & Stutman, O. *J. Immun.* **121**, 1827–1835 (1978).
6. Kumar, V., Luevano, E. & Bennett, A. *J. exp. Med.* **150**, 531–547 (1979).
7. Burton, R. C. in *Natural Cell-mediated Immunity Against Tumors*, 19–35 (Academic, New York, 1980).
8. Stutman, O., Feo Figarella, E., Paige, C. J. & Lattime, E. C. in *Natural Cell-mediated Immunity Against Tumors*, 187–229 (Academic, New York, 1980).
9. Stutman, O. in *Natural Cell-mediated Immunity Against Tumors*, 231–240 (Academic, New York, 1980).
10. Stutman, O., Dien, P., Wisun, R. & Lattime, E. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2895–2898 (1980).
11. Roder, J. & Duwe, A. *Nature* **278**, 451–453 (1979).
12. Roder, J. *J. Immun.* **123**, 2168–2173 (1979).
13. Roder, J., Lohmann-Matthes, M. L., Domzig, W. & Wigzell, H. *J. Immun.* **123**, 2174–2181 (1979).
14. Padgett, G. A., Holland, J. M., Prieur, D. J., Davis, W. C. & Gorham, J. R. in *Animal Models for Biomedical Research* Vol. 3, 1–12 (National Academy of Sciences, Washington DC, 1970).
15. Haliotis, T. et al. *J. exp. Med.* **151**, 1039–1048 (1980).
16. Stutman, O. & Cuttito, M. J. in *Natural Cell-mediated Immunity Against Tumors*, 431–442 (Academic, New York, 1980).
17. Lane, P. W. & Murphy, E. D. *Genetics* **72**, 451–460 (1972).
18. Brandt, E. J. & Swank, R. T. *Am. J. Path.* **82**, 573–588 (1976).
19. Talmadge, J. E., Meyers, K. M., Prieur, D. J. & Starkey, J. R. *Nature* **284**, 622–624 (1980).
20. Karre, K., Klein, G. O., Kiessling, R., Klein, G. & Roder, J. C. *Nature* **284**, 624–626 (1980).
21. Salomon, J. C., Creau-Goldberg, N. & Lynch, N. R. *Cancer Immun. Immunother.* **8**, 67–70 (1980).
22. Stutman, O. in *The Nude Mouse in Experimental and Clinical Research*, 411–435 (Academic, New York, 1978).
23. Dent, P. B., Fish, L. W., White, J. F. & Good, R. A. *Lab. Invest.* **15**, 1634–1642 (1966).
24. Murphy, E. D. in *Biology of the Laboratory Mouse*, 521–562 (McGraw-Hill, New York, 1966).
25. Heston, W. E. *J. natn. Cancer Inst.* **3**, 79–82 (1942).
26. Strong, L. C. *Yale J. Biol. Med.* **25**, 34–43 (1952).
27. Stutman, O. *Adv. Cancer Res.* **22**, 261–422 (1975).
28. Boraschi, D. & Meltzer, M. E. *J. Immun.* **124**, 1050–1053 (1980).
29. Burnet, F. M. *Prog. exp. Tumour Res.* **13**, 1–27 (1970).
30. Stutman, O. in *The Handbook of Cancer Immunology* Vol. 7, 1–25 (ed. Waters, H.) (Garland STPM, New York, 1980).
31. Durdik, J. M., Beck, B. N. & Henney, C. S. in *Natural Cell-mediated Immunity Against Tumors*, 37–46 (Academic, New York, 1980).
32. Brunda, M. J., Holden, H. T. & Heberman, R. B. in *Natural Cell-mediated Immunity Against Tumors*, 411–415 (Academic, New York, 1980).

33. Chun, M., Pasanen, V., Hammerling, U., Hammerling, G. F. & Hoffman, M. K. *J. exp. Med.* **150**, 426–431 (1979).
34. Young, W. W., Hakomori, S. I., Durkik, J. M. & Henney, C. S. *J. Immun.* **124**, 199–201 (1980).
35. Mahoney, K. H., Morse, S. S. & Morahan, P. S. *Cancer Res.* **40**, 3934–3939 (1980).
36. Riccardi, C., Santoni, A., Marlozzari, T. & Herberman, R. B. in *Natural Cell-mediated Cytotoxicity Against Tumors*, 1121–1139 (Academic, New York, 1980).
37. Mattes, M. J., Sharrow, S. O., Herberman, R. B. & Holden, H. T. *J. Immun.* **123**, 2851–2890 (1979).
38. Koo, G. C., Jacobson, J. B., Hammerling, G. J. & Hammerling, U. *J. Immun.* **125**, 1003–1006 (1980).
39. Stutman, O., Lattime, E. C. & Feo Figarella, E. *Fedn Proc.* (in the press).

Idiotypic regulation by isologous monoclonal anti-idiotope antibodies

M. Reth, G. Kelsoe & K. Rajewsky

Institute for Genetics, University of Cologne, Weyertal 121, D-5000 Köln 41, FRG

The network hypothesis¹ proposes that internal antibody interactions are crucial in the regulation of the immune system. Thus, injection of minute amounts of anti-idiotypic antibodies (antibodies specific for a particular set of antigenic determinants (idiotopes) associated with the variable region of a particular antibody or set of antibodies) can enhance or suppress the production of the corresponding set of idiotopes (the idiotype) in subsequent antibody responses of the animal^{2–5}. We have recently shown that idiotope expression can be enhanced and suppressed by monoclonal anti-idiotope antibodies of the IgG1 class in concentrations approaching those of natural antibodies⁶. We now demonstrate that these regulatory effects are not due to differences in the constant parts of the injected antibody and antibodies of the IgG1 class in the recipient. Also, an IgG2a anti-idiotope antibody raised in C57BL/6 mice suppresses the expression of the corresponding idiotope in animals of the same strain when minute amounts are injected 2–8 weeks before the test immunization with antigen. These results indicate that antibody-mediated network interactions are physiologically important and suggest that the class of isologous anti-idiotypic antibodies may influence their regulatory properties. Finally, the enhancement and suppression of idiotope expression by IgG1 anti-idiotopes and the suppression by IgG2a is almost exclusively at the level of IgG but not IgM antibodies.

We developed previously, by the Köhler–Milstein cell hybridization technique⁷, a series of monoclonal antibodies of mouse origin each of which recognizes a particular idiotope characteristic for a subset of anti-NP (NP = 4-hydroxy-3-nitrophenyl-acetyl) antibodies of C57BL/6 mice⁸. When IgG1 antibodies of this type were injected into C57BL/6 mice and the animals immunized with an NP-carrier conjugate 1–8 weeks later, the expression of the corresponding idiotope in the anti-NP antibody population was either enhanced or suppressed, depending on the dose of anti-idiotypic antibody: nanogramme amounts led to enhancement, microgramme amounts to suppression (ref. 6 and unpublished observations).

The antibodies used in these experiments originated from CBA mice and thus differed in allotype from IgG1 antibodies of the C57BL/6 strain. The experiment was therefore repeated in (C57BL/6 × CBA)F₁ animals for which the anti-idiotope antibody would be 'self', at least with respect to its constant part. Groups of five animals were injected with 10 ng or 10 µg antibody Ac38 and immunized with NP₈-CG (CG = chicken γ-globulin) 4 or 6 weeks later. The experimental conditions were exactly as described in our previous paper⁶, except that the anti-idiotope antibodies were diluted in phosphate-buffered saline containing 5 µg mouse serum albumin per ml. The animals were bled 12 days after immunization with antigen and the sera were titrated for λ₁-bearing anti-NP antibodies and the

Ac38 idiotope (the idiotope recognized by antibody Ac38). This idiotope is expressed within the λ-bearing antibody population which represents a major portion of the total primary anti-NP response in mice expressing the *Igh^b* haplotype^{9,10}. The results of the experiment are depicted in Fig. 1.

It is clear that 10 ng of antibody Ac38 enhance and 10 µg of the antibody suppress the expression of the Ac38 idiotope in the primary anti-NP response. There is no effect on the total amount of λ₁-bearing anti-NP antibodies. These results are strictly analogous to those previously obtained in C57BL/6 mice⁶ and thus demonstrate that allotype disparity is not responsible for the regulatory effect of these anti-idiotope antibodies.

Sera from one of the two experiments depicted in Fig. 1 were analysed for the class distribution of idiotope-bearing and total anti-NP antibodies. The sera from the control groups were not available for analysis, as they had been used up in absorption experiments in which it was shown that elimination of NP-binding molecules from the sera also led to the elimination of more than 90% of idiotope-bearing molecules. We therefore included in the analysis sera from our previous experiments⁶, in which antibody Ac38 had been injected into C57BL/6 mice instead of (CBA × C57BL/6)F₁ animals. The results are shown in Fig. 2.

The sera from the F₁ experiment were titrated for IgM, IgG1, IgG2a and IgG2b (Fig. 2a). IgG2a anti-NP antibodies were not detectable and IgG1 was the predominant class. The level of antibodies bearing the Ac38 idiotope was high in IgG in the 10-ng group and reduced to background in the 10-µg group.

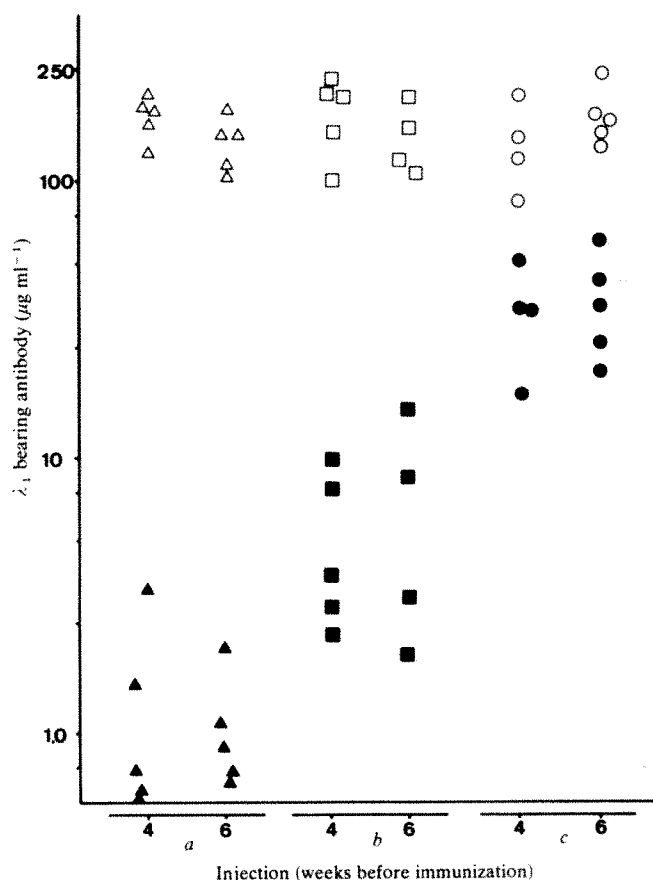


Fig. 1 Serum antibody levels from (CBA × C57BL/6)F₁ mice injected with: a, 10 µg Ac38 (△, ▲); b, diluent (□, ■); and c, 10 ng Ac38 (○, ●) 4 or 6 weeks before immunization with 100 µg NP₈-CG. Mice were bled 12 days after immunization and individual serum samples collected for analysis. Individual values from the results of a solid-phase radioimmunoassay are plotted as µg ml⁻¹ of λ₁-bearing antibody⁶. Open symbols represent the total amount of NP-specific antibody, closed symbols represent that fraction of antibody bearing the Ac38 idiotope.

In contrast, at the level of IgM, both groups exhibited similar levels of Ac38-positive anti-NP antibodies. The notion that the regulation of idiotype expression by antibody Ac38 operates mainly at the level of the IgG response is further supported by the data of the experiment with C57BL/6 mice (Fig. 2*b*). In this case the sera were analysed for idiotype expression in IgM and IgG1 as the major IgG component. Obviously enhancement of Ac38 expression appears mainly in IgG1 and not at all in IgM; suppression is complete in IgG1 and only partial in IgM.

One might still argue that the experiments documented in Figs 1 and 2 do not reflect a physiological process in the immune system because antibody Ac38 was not induced in the strain expressing the Ac38 idiotype (C57BL/6), although this argument does not seem to make much sense in view of the extreme diversity of the antibody system. Nevertheless, we have started to produce, in the same idiotypic system, monoclonal anti-idiotype antibodies of C57BL/6 origin. For this purpose C57BL/6 mice were sensitized with a conjugate of keyhole limpet haemocyanin and the anti-NP antibody B1-8, a monoclonal IgM of C57BL/6 origin—the same protein against which our previous anti-idiotype antibodies had been raised⁸. The animals were boosted with the same conjugate 2 weeks later and their spleen cells fused with the myeloma line X63.Ag8-6.5.3 (ref. 11) 3 days after boosting. All technical details of immunizations, the cell fusion technique⁷, the detection, cloning and propagation of cells secreting monoclonal anti-idiotype antibodies, the characterization of these antibodies and their purification have been described previously⁸. A cell line designated A6.24 was isolated, cloned and propagated in (C57BL/6 × BALB/c)F₁ animals; it secreted large amounts of a

monoclonal anti-B1-8 antibody. The antibody was purified from ascites fluid by ammonium sulphate precipitation and subsequent chromatography on DEAE cellulose (DE52, Whatman) as described previously⁸. The purified material showed a single band in microzone electrophoresis, and on SDS-polyacrylamide electrophoresis¹² in reducing conditions, two main bands appeared in the position of immunoglobulin L and H chains. It typed as an IgG2a in an enzyme-linked immunosorbent assay¹¹. In radioactive binding assays antibody A6.24 reacted with the B1-8 protein and its Fab and Fv fragment⁸ and a fraction of C57BL/6 anti-NP serum antibodies (see below), but not with any of a series of myeloma proteins or, at a detectable level, with immunoglobulins from normal sera. The reaction of antibody A6.24 with B1-8 could be specifically blocked by an excess of NIP-BSA (4-hydroxy-5-iodo-3-nitrophenyl-bovine serum albumin) but not of NIP-caproic acid. Antibody A6.24 thus behaves as a typical anti-idiotype antibody detecting a group 2 idiotype in our previous nomenclature⁸. This idiotype is distinct from the idiotype defined by antibody Ac38, as antibodies A6.24 and Ac38 exhibit different patterns of reactivity with a series of NP-binding monoclonal antibodies (data not shown).

Antibody A6.24 is also distinguished from the previously used antibody Ac38 by its origin (C57BL/6) and by its class: it types as an IgG2a whereas Ac38 is an IgG1. How does injection of antibody A6.24 into C57BL/6 mice influence the expression of the corresponding idiotype (which we call idiotype A6.24) in the anti-NP response? Groups of age- and sex-matched C57BL/6 mice received 33 ng or 10 µg A6.24 intraperitoneally. Controls were injected with diluent alone. 2–8 weeks later the animals were immunized with NP₈-CG and bled 12 days later. The protocol of the experiment thus followed exactly that of our previous, analogous experiments (see above and ref. 6). The sera were titrated for total λ₁-bearing anti-NP antibodies and for idiotype A6.24. The results of these titrations appear in Fig. 3.

It can be seen that, as in the experiments with Ac38, 10 µg A6.24 profoundly suppresses the expression of the corresponding idiotype, leaving the total anti-NP response untouched. Suppression is almost complete throughout the entire experiment. However, at the 33-ng dose which would have led to maximal enhancement in the case of Ac38⁶, antibody A6.24 again produces suppression, although less pronounced than at the 10-µg dose. As previously⁶, there is an indication of oscillations of idiotype expression. The class distribution of the anti-NP antibodies in the sera from various groups of this experiment was also determined (data not shown). It was essentially the same as that in the experiment depicted in Fig. 2. Only idiotype-bearing IgG but not IgM antibodies were significantly suppressed.

Our data show that striking modulation of idiotype expression in an immune response can be achieved by the prior administration of an isologous, monoclonal anti-idiotype antibody. Thus, the capacity to recognize, respond against and regulate the expression of a given idiotype determinant exists within a single mouse strain. This strongly suggests that antibody-mediated network regulation is an important physiological process as discussed in detail elsewhere⁶. Our data also suggest that the effector function of isologous anti-idiotype antibodies in network regulation may be influenced by the heavy-chain constant region. Antibody Ac38, an IgG1, enhances and suppresses, depending on the dose. The same has been found for another IgG1 anti-idiotype, Ac146 (ref. 13). Antibody A6.24, an IgG2a, may only suppress. This is reminiscent of the previous finding that guinea pig IgG1 anti-idiotypic antibodies enhance idiotype expression in the mouse, whereas IgG2 has suppressive effects^{2,3}. At least in terms of their complement-fixing properties, IgG1 and IgG2 in mice and guinea pigs are similar^{14,15}. Clearly, however, our experiments can not formally assign distinct functions in network control to the constant regions of γ1 and γ2a heavy chains. It is known, for example, that different allelic forms of the γ2a chain may differ drastically from each other in primary structure (P. Schreier and

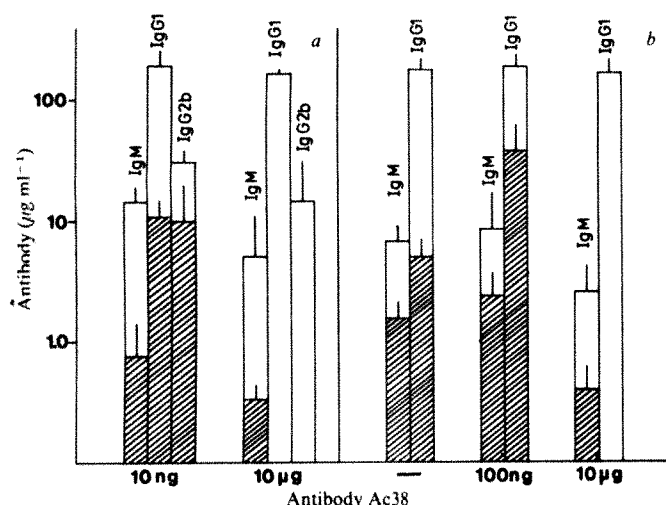


Fig. 2 Immunoglobulin class distribution of total and idiotype Ac38-bearing anti-NP antibodies in the sera of groups of mice injected with different doses of anti-idiotype antibody (Ac38) and subsequently immunized with 100 µg NP₈-CG. Each column represents the geometric mean ($\pm$ s.d.) of the amounts of anti-NP antibody of a particular immunoglobulin class expressed within a group of mice. That fraction of antibodies bearing the Ac38 idiotype is given by the shaded portion of each column. *a*, (CBA × C57BL/6)F₁ mice injected with 10 ng or 10 µg of Ac38 6 weeks before immunization with NP-CG. *b*, C57BL/6 mice injected with diluent, 100 ng Ac38, or 10 µg Ac38. Based on earlier results⁶, mice injected with 10 µg or 100 ng Ac38 were immunized with NP-CG 4 or 8 weeks later, respectively. Antibody concentrations were determined in a solid-phase radioimmunoassay⁶. To determine total anti-NP antibody, hapten-binding immunoglobulin was detected with ¹²⁵I-labelled, immunoglobulin class-specific antibodies (described in ref. 16). Those antibodies bearing the Ac38 idiotype were detected in a similar manner with the exception that the solid-phase coat was the F(ab)₂ fragment of antibody Ac38 rather than a hapten-protein conjugate. Absorption experiments showed that virtually all (>90%) idiotype Ac38-positive antibodies bound the NP hapten.

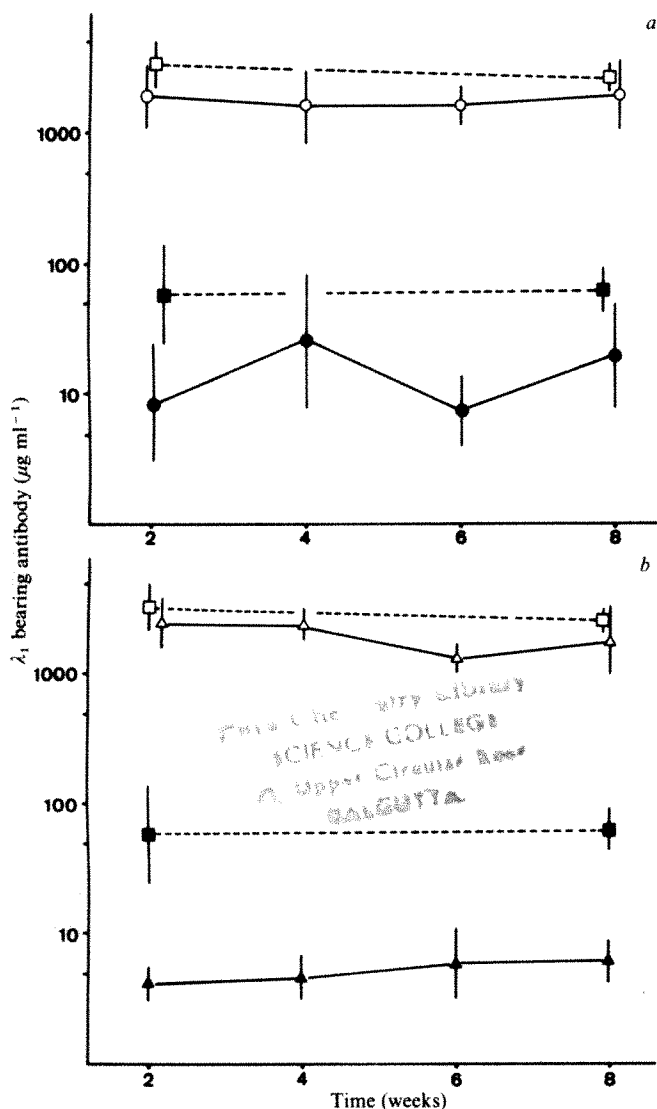


Fig. 3 Idiotope suppression induced by monoclonal anti-idiotope antibody A6.24. Groups of six mice were injected intraperitoneally with *a*, 33 ng (○, ●; unbroken lines) or *b*, 10 µg (△, ▲; unbroken lines) of antibody A6.24 and immunized 2–8 weeks later with NP-CG. Control mice (□, ■; broken lines) received diluent instead of anti-idiotope antibody. 12 days after the administration of antigen the animals were bled and the sera of individual mice analysed in a solid-phase radioimmunoassay⁶ for total, λ_1 -bearing anti-NP antibodies (open symbols) and λ_1 -bearing anti-NP antibodies expressing the A6.24 idiotope. The geometric means (with s.d.) of the antibody concentrations in the sera are plotted against the time between injection of anti-idiotope antibody and antigen.

A. Bothwell, personal communication). A definitive assignment will require the testing of a number of isologous anti-idiotope antibodies of various isotypes and (more importantly) the isolation of pairs of anti-idiotope antibodies with identical variable regions but belonging to different immunoglobulin classes, for example class-switch variants. A technique for the isolation of such pairs of antibodies is available¹⁶. Finally, it seems that the target of regulation by our anti-idiotope antibodies is mainly the IgG and not the IgM antibody response. It will be important to investigate the basis of this finding as regulatory studies with conventional isologous anti-idiotypic antisera suggest that in certain conditions, IgM and IgE responses can also be affected by idiotypic control^{4,17}.

This investigation was supported by the Deutsche Forschungsgemeinschaft through SFB 74. G.K. is a recipient of a Fellowship from the US NSF. We thank Ms Gertrud von Hesberg for technical assistance.

Received 26 September 1980; accepted 20 January 1981.

1. Jerne, N. K. *Annls. Immun. Inst. Pasteur, Paris* **125C** 373 (1974).
2. Eichmann, K. *Eur. J. Immun.* **4**, 296 (1974).
3. Eichmann, K. & Rajewsky, K. *Eur. J. Immun.* **5**, 661 (1975).
4. Köhler, H. *Transpl. Rev.* **27**, 24 (1975).
5. Julius, M. H., Augustin, A. A. & Cosenza, H. in *Regulatory Genetics of the Immune System* (eds Sercarz, E. E., Herzenberg, L. A. & Fox, C. F.) Vol. 6, 179 (Academic, New York, 1977).
6. Kelsoe, G., Reth, M. & Rajewsky, K. *Immun. Rev.* **52**, 75 (1980).
7. Köhler, G. & Milstein, C. *Nature* **256**, 495 (1975).
8. Reth, M., Imanishi-Kari, T. & Rajewsky, K. *Eur. J. Immun.* **9**, 1004 (1979).
9. Imanishi-Kari, T., Rajnavölgyi, E., Takemori, T., Jack, R. S. & Rajewsky, K. *Eur. J. Immun.* **9**, 324 (1979).
10. Karjalainen, K. *Eur. J. Immun.* **10**, 132 (1980).
11. Kearney, J. F., Radbruch, A., Liesegang, B. & Rajewsky, K. *J. Immun.* **123**, 1548 (1979).
12. Laemmli, K. K. *Nature* **227**, 680 (1970).
13. Kelsoe, G., Reth, M. & Rajewsky, K. *Eur. J. Immun.* (in the press).
14. Klaus, G. G. B., Pepys, M. B., Katajima, K. & Askonas, B. A. *Immunology* **38**, 687 (1979).
15. Bloch, K. J., Kourilsky, F. M., Ovary, Z. & Benacerraf, B. *J. exp. Med.* **117**, 965 (1963).
16. Radbruch, A., Liesegang, B. & Rajewsky, K. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2909 (1980).
17. Blaser, K., Nakagawa, T. & de Weck, A. J. *Immun.* **125**, 24 (1980).

Control of enzyme secretion by non-cholinergic, non-adrenergic nerves in guinea pig pancreas

G. T. Pearson, J. S. Davison, R. C. Collins & O. H. Petersen*

Department of Physiology, University of Dundee, Dundee DD1 4HN, UK

Depolarization of pancreatic cells by exposure to high potassium solutions is associated with release of amylase. In the guinea pig, but not the mouse or cat, this Ca-dependent amylase secretion is resistant to atropine blockade^{1–4}, thus Scheele and Haymovits^{3,4} concluded that the enzyme secretion evoked by K depolarization does not involve release of transmitter from intrapancreatic nerves but is a consequence of Ca uptake into acinar cells mediated by the membrane depolarization. This hypothesis is inconsistent with current concepts of stimulus–secretion coupling in electrically non-excitable cells⁵. The observation of Scheele and Haymovits could, however, also be explained by the release of a non-cholinergic, secretomotor transmitter as a consequence of the depolarization of intrapancreatic nerves. By adapting the technique of electrical field stimulation of isolated pancreatic segments^{6,7} to our studies of amylase secretion, we have now been able to demonstrate both cholinergic and non-cholinergic, non-adrenergic secretomotor nerves in the guinea pig pancreas. Excitation of the non-cholinergic nerves stimulates amylase secretion by a different intracellular coupling mechanism from that activated by cholinergic nerves or by peptides belonging to the cholecystikinin, gastrin or bombesin families⁸.

All experiments were done on isolated segments of guinea pig pancreas superfused with oxygenated physiological salt solutions at 37 °C (ref. 7). Amylase secretion from the pancreas was measured using an on-line automated fluorometric method⁹, intracellular recordings of acinar transmembrane potentials were obtained using glass microelectrodes⁷, and Ca efflux was monitored by measuring the release of ⁴⁵Ca from prelabelled tissue¹⁰.

Electrical field stimulation (FS) has recently been shown to be an effective method of exciting intrinsic nerves in the pancreas^{6,7}. Using Perspex flow cells incorporating silver electrodes, we have studied the effects of FS on membrane potentials, amylase secretion and ⁴⁵Ca²⁺ efflux. In preliminary experiments we confirmed the results of Scheele and Haymovits^{3,4} that an elevated extracellular K concentration (75 mM) stimulated amylase secretion even in the presence of atropine (10^{–5} M). This protocol, however, is unhelpful as it is impossible to distinguish between indirect effects mediated by depolarizing nerve endings and direct effects due to depolariza-

* To whom correspondence should be addressed.

tion of the acinar cells. The nerve-blocking drug tetrodotoxin (TTX) (10^{-6} M) did not abolish the effects, as also noted by Scheele and Haymovits⁴, but K depolarization of nerve endings might of course evoke transmitter release even in the absence of action potential formation. We decided, therefore, to investigate more directly the possible existence of non-cholinergic nerves.

Figure 1a shows FS-evoked amylase secretion from a tissue superfused with normal physiological solution. Addition of the muscarinic antagonist atropine (10^{-5} M), in a concentration sufficient to abolish maximum response to exogenous acetylcholine (ACh) stimulation (10^{-4} M), reduced, but did not abolish, the FS response. Such an atropine-resistant amylase

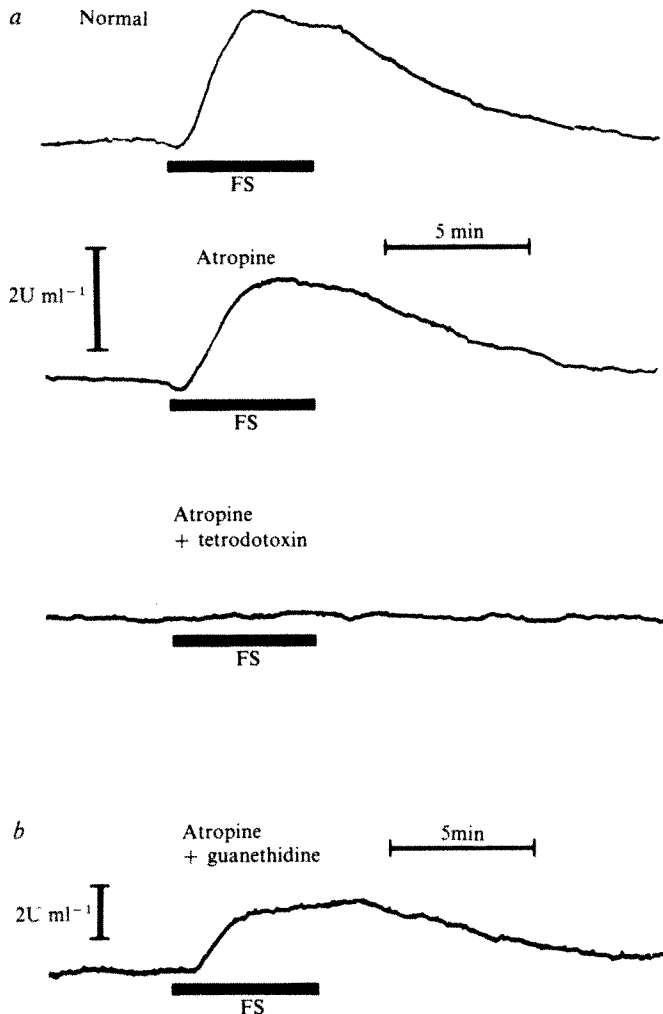


Fig. 1 Effects of electrical field stimulation (FS) on amylase release from superfused guinea pig pancreatic fragments. The figure shows two separate experiments (a and b) in which amylase in the effluent from the tissue chamber (1 ml vol, flow rate 1 ml min⁻¹) was continuously assayed by an automated fluorometric, on-line method⁹. FS, for the duration indicated, was achieved by connecting silver electrodes in the chamber to an isolated stimulator. Stimulation parameters were the same throughout: 20 Hz frequency, 50 V amplitude, 1 ms pulse width for 5 min. (FS using these parameters, even for periods up to 20 min, did not cause any change in the pH of the Krebs solution flowing through the tissue chamber.) The three panels in a are consecutive traces from the same experiment showing the effects of FS on the amylase concentration in the tissue chamber effluent during exposure of the tissue to a normal physiological salt solution, in the presence of atropine (10^{-5} M) and in the presence of both atropine (10^{-5} M) and TTX (10^{-6} M); b shows the FS response in the presence of both atropine (10^{-5} M) and guanethidine (10^{-5} M). All drugs applied were present in the superfusing solution for a minimum period of 20 min before FS. Note the different calibrations (2 units of amylase ml⁻¹) in a and b.

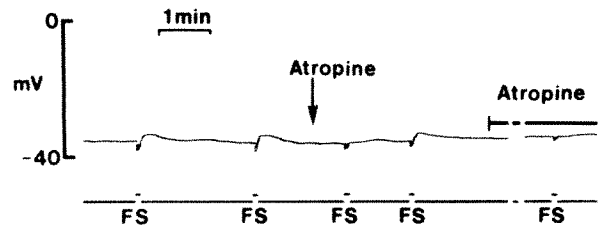


Fig. 2 Recording of guinea pig acinar cell membrane potential. An isolated segment of pancreas was secured to a Perspex platform and superfused with oxygenated physiological salt solution at 37 °C. A single microelectrode filled with 3 M KCl was inserted in an acinar cell to record a resting membrane potential of -35 mV. FS, as described in detail elsewhere⁷, was achieved using a pair of silver wire electrodes placed in contact with the tissue on the Perspex platform. These were connected to an isolated stimulator and the stimulation parameters throughout the experiment were 40 Hz, 30 V, 1 ms for 5 s. The effect of FS on the cell membrane potential is seen both in the absence and presence of atropine (10^{-5} M). At the arrow the solution flowing into the bath contained 10^{-5} M atropine for only ~10 s, thus the bath concentration of atropine is unlikely to have attained the 10^{-5} M concentration. Nevertheless, as seen in the trace, the FS effect was abolished. The response, however, rapidly returned and subsequent, sustained exposure to 10^{-5} M atropine (bar) again abolished the effect of FS. The break in the recording represents a 2-min interval.

secretion has been observed in 16 separate guinea pig preparations. Subsequent addition of TTX in a concentration that does not affect ACh or caerulein responses in the pancreas⁷ completely abolished the effect of FS. Abolition of the atropine-resistant FS response by TTX was observed in all six preparations tested, indicating that FS causes amylase release by initiating action potentials in intrapancreatic nerves. The adrenergic neurone-blocking drug, guanethidine (10^{-5} M), had no effect on the nerve-mediated amylase secretion in the five preparations tested (Fig. 1b), suggesting that the non-cholinergic response is not mediated by stimulation of adrenergic nerves. Furthermore, we have been unable to demonstrate any effect of noradrenaline (6×10^{-5} M) on amylase secretion. Although most experiments measuring nerve-mediated amylase secretion have been done using FS at 20 Hz frequency, effects have been observed at frequencies as low as 5 Hz. The mean value ($\pm$ s.e.m.) for FS-evoked amylase release at 20 Hz in the presence of atropine was $0.45(\pm 0.06)$ U per ml per 100 mg ($n=10$), corresponding to 20% of the mean value for the maximal secretion evoked by exogenous ACh (10^{-4} M). We conclude that the amylase secretion evoked by FS in the presence of atropine (Fig. 1) is mediated by excitation of intrapancreatic nerves, releasing a transmitter substance which activates the pancreatic acinar cells. The transmitter is not ACh and there is no indication that it is a catecholamine.

To test the functional innervation of individual acinar units⁷, membrane potential changes in response to FS were studied. Figure 2 shows examples of membrane depolarizations evoked by FS (40 Hz). Of 84 units studied, only 23 responded to FS, in marked contrast to the 100% effectiveness in the mouse⁷ and rat pancreas¹¹. The FS-evoked depolarizations, shown in Fig. 2, are always abolished by atropine (five preparations). It seems, therefore, that the transmitter released by excitation of the non-cholinergic, non-adrenergic nerves is unable to evoke membrane depolarization in spite of being a potent secretagogue.

The pancreatic secretagogues evoking enzyme secretion mediated by changes in cellular calcium metabolism (for example, ACh, cholecystokinin and bombesin) also cause acinar cell depolarization, while the secretagogue secretin (or vasoactive intestinal polypeptide, VIP) has no effect on ⁴⁵Ca fluxes or membrane electrophysiological parameters^{8,12,13}. We therefore tested the ability of FS to alter calcium metabolism in the guinea pig pancreas.

In the absence of atropine, FS caused a clear increase in fractional ^{45}Ca efflux from preloaded pancreatic segments in two separate experiments (Fig. 3). In the presence of atropine, however, the same type of FS (20 Hz, 40 V, 1 ms) failed to evoke any measurable effect on ^{45}Ca efflux in the four different preparations tested, although all of these responded with a marked increase in ^{45}Ca outflux when subsequently challenged with caerulein (1.5×10^{-9} M, Fig. 3).

The nerve-mediated secretion obtained in the presence of atropine (Fig. 1) is clearly not associated with membrane depolarization (Fig. 2) or with an increase in ^{45}Ca efflux (Fig. 3). The as yet unidentified transmitter substance is therefore unlikely to be a peptide belonging to the cholecystokinin-gastrin-caerulein, bombesin or substance P families because all secretagogues studied from these groups characteristically evoke marked Ca outflux and membrane electrical changes^{8,12,13}.

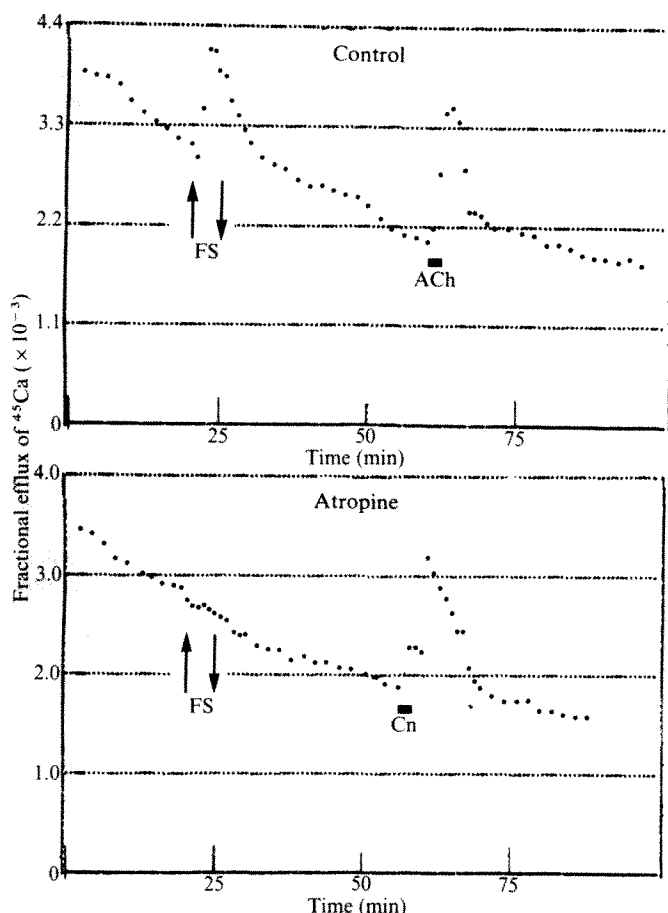


Fig. 3 The effect of FS and the secretagogues caerulein (Cn) and acetylcholine (ACh) on the fractional efflux of ^{45}Ca from preloaded superfused fragments of guinea pig pancreas. The pancreatic segments were loaded for 60 min in 2 ml of physiological salt solution containing $10 \mu\text{Ci } ^{45}\text{Ca}$ and maintained at 37°C . The procedure was similar to that described elsewhere¹⁰. After loading, the segments were transferred to a flow cell. In each experiment the tissue was superfused for 150 min before collection of effluent samples, which were subsequently analysed in a scintillation counter. Values for concentration of radioactivity in the samples taken at different time points together with the value for the radioactivity remaining in the tissue at the end of the experiment were processed by computer (MINC-11, Digital). Values were obtained for fractional efflux ($\Delta x / \Delta t x_i \text{ min}^{-1}$, where x represents c.p.m. ^{45}Ca released in the time interval Δt and x_i the tissue ^{45}Ca content at the midpoint of interval Δt) as a function of time. FS was carried out in exactly the same way as described for the amylase studies. Stimulation parameters were similar in each experiment: 20 Hz, 40 V, 1 ms for 5 min. Cn (1.5×10^{-9} M) and ACh (10^{-5} M) were added to the superfusing solution for the periods indicated by the bars. Atropine (10^{-5} M) was present throughout the experiment shown in the lower panel. Upper panel, normal preparation.

The present study has demonstrated, for the first time, the existence of non-cholinergic, non-adrenergic nerves controlling pancreatic acinar enzyme secretion. Its importance lies in the widespread distribution of non-cholinergic, non-adrenergic nerves throughout the peripheral autonomic nervous system^{14,15}. The transmitter released from the non-cholinergic, non-adrenergic nerves in our experiments might be VIP or a related peptide, as VIP-immunoreactive nerves have been demonstrated in the pancreas^{16,17}. VIP is known to stimulate amylase secretion by a mechanism not involving changes in cellular Ca metabolism^{8,13}.

Our results also resolve an important controversy in the discussion of pancreatic stimulus-secretion coupling⁸ because the Ca-dependent depolarization-evoked enzyme secretion observed in the presence of atropine^{3,4} can now most conveniently be explained as being due to Ca-dependent secretion of a neurotransmitter that interestingly does not seem to activate secretion through changes in cellular calcium metabolism.

This work was supported by grants from the MRC to O.H.P. We thank Drs I. Findlay and A. R. Chipperfield for advice and Dr M. S. Daoud and Miss K. Watson for technical assistance.

Received 5 November 1980; accepted 2 February 1981.

- Argent, B. E., Case, R. M. & Scratcherd, T. *J. Physiol., Lond.* **216**, 611-624 (1971).
- Poulsen, J. H. & Williams, J. A. *Pflügers Arch. ges. Physiol.* **370**, 173-177 (1977).
- Scheele, G. & Haymovits, A. *Ann. N.Y. Acad. Sci.* **307**, 648-652 (1978).
- Scheele, G. & Haymovits, A. *J. Biol. Chem.* **255**, 4918-4927 (1980).
- Petersen, O. H. *Physiol. Rev.* **56**, 535-577 (1976).
- Nishiyama, A., Katoh, K., Saitoh, S. & Wakui, M. *Membr. Biochem.* **3**, 49-66 (1980).
- Davison, J. S., Pearson, G. T. & Petersen, O. H. *J. Physiol., Lond.* **301**, 295-305 (1980).
- Schulz, I. & Stolze, H. H. *Rev. Physiol.* **42**, 127-156 (1980).
- Matthews, E. K., Petersen, O. H. & Williams, J. A. *Analyt. Biochem.* **58**, 155-160 (1974).
- Matthews, E. K., Petersen, O. H. & Williams, J. A. *J. Physiol., Lond.* **234**, 689-701 (1973).
- Davison, J. S., Pearson, G. T. & Petersen, O. H. in *Biology of Normal and Cancerous Exocrine Pancreatic Cells* (eds Ribet, A., Pradayrol, L. & Susini, C.) 83-88 (Elsevier, Amsterdam, 1980).
- Petersen, O. H. & Iwatsuki, N. *Ann. N.Y. Acad. Sci.* **307**, 599-617 (1978).
- Gardner, J. D. A. *Rev. Physiol.* **41**, 55-66 (1979).
- Burnstock, G. *Pharmac. Rev.* **24**, 509-581 (1972).
- Burnstock, G. in *Physiological and Regulatory Functions of Adenosine and Adenine Nucleotides* (eds Baer, H. P. & Drummond, G. I.) 3-32 (Raven, New York, 1979).
- Larsson, L.-I., Fahrenkrug, J., Holst, J. J. & Schaffalitzky de Muckadell, O. B. *Life Sci.* **22**, 773-780 (1978).
- Sundler, F., Alumets, J., Håkanson, R., Fahrenkrug, J. & Schaffalitzky de Muckadell, O. B. *Histochemistry* **55**, 173-176 (1978).

Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts

Chiaho Shih, L. C. Padhy, Mark Murray & Robert A. Weinberg

Department of Biology and Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

We have previously¹ demonstrated that DNA of mouse fibroblasts transformed by 3-methylcholanthrene (3-MC) induced foci of transformed cells when applied to monolayer cultures of NIH3T3 cells, which indicates that at least a part of this phenotype is encoded in DNA sequences. However, our conclusions were confined to the effects of DNAs of 3-MC-transformed mouse fibroblasts on recipient NIH3T3 cells, also of mouse fibroblast origin. To elucidate this phenomenon further, we have prepared DNAs from a series of mouse and non-mouse tumour lines of non-fibroblastic origin and investigated whether tumour transforming genes can act across tissue and species barriers to transform NIH3T3 cells. We find that DNAs obtained from human, rabbit and mouse bladder carcinoma lines, a lung carcinoma line and rat neuroblastoma and mouse glioma lines, are able to induce transformation of NIH3T3 cells on transfection.

DNAs prepared from tumours or tumour cell lines were transfected using the calcium phosphate transfection procedure described elsewhere^{1,2}. Each transfected culture was passaged into 12 subcultures and foci of transformants were scored 14–20 days later¹; only foci containing very refractile cells were counted. Because of occasional ambiguities in resolving true transformants from spontaneous overgrowths of the monolayer culture, critical transfections were scored using a double-blind protocol. We observed induction of transformed foci when transfected DNAs were prepared from two classes of tumour cells: neuroblastoma/glioma (Table 1) and carcinoma (Table 2).

Most of the DNAs used in these experiments were prepared from tumour cell lines which had been derived from tumour explants. Each tumour cell line that gave positive results had been shown to possess differentiation markers characteristic of the tumour type from which the cell line had been derived^{3–8}. For example, the neuroblastoma lines had been shown to have excitable membranes, neurite extensions and neurotransmitter enzymes³; tumours induced by cell lines of carcinoma origin possessed keratin and characteristic acinar structures⁵.

As seen in Table 1, foci were induced by the DNAs of one glioma and three neuroblastoma cell lines. The latter were derived from tumours induced in rats *in utero* by exposure to ethylnitrosourea³, and the glioma cell was prepared from a tumour induced by direct implantation of a 3-MC precipitate into the brain of a newborn mouse⁴. The foci produced by these transfections were grown up and their DNA prepared for a second cycle of transfection. The DNAs of these primary foci were all able to induce secondary foci, which demonstrates the serial transmissibility of these transformation-inducing sequences.

Transfection of the DNAs of three different bladder carcinoma cell lines also yielded foci after application to NIH3T3 monolayer cultures as did transfection of DNAs of the mouse Lewis lung carcinoma tumour (Table 2). Again, DNAs prepared from cultures of these primary transfectants were able to induce foci after a second cycle of transfection. The DNAs of a large series of other carcinoma cell lines yielded few or no foci on transfection (Table 2).

When the cells of foci resulting from transfections induced by neuroblastoma, glioma and carcinoma DNAs were grown into mass cultures and seeded subcutaneously into newborn NFS mice, rapidly growing fibrosarcoma-like tumours were observed

after 2 weeks with almost 100% incidence (Tables 1 and 2). Thus, the transfected DNA induced foci of cells which were highly tumorigenic. In contrast, normal, untransfected NIH3T3 cells resulted in a low incidence of tumours (1/7, 1/10) which only appeared 4 weeks after subcutaneous injection into the newborn mice (Tables 1, 2).

All attempts at rescue of C-type transforming retroviruses from the transfected cells were unsuccessful. We also attempted rescue of virus from cells transformed by Moloney sarcoma virus and observed the release of high titres of focus-forming virus. We conclude that transmissible C-type viruses are not responsible for the phenomenon of focus induction in donor or recipient cells.

Human donor DNA was used in one of these transfections to obtain direct biochemical proof that the transfected cells had taken up and stabilized human DNA. This depended on the fact that the few cells which acquire a selected marker gene also acquire a large number of other donor DNA fragments which enter the cells during the same transfection event⁹. In contrast, other cells in the culture which have not acquired the selected marker gene do not show this array of co-transfected donor sequences⁹.

In the present study, cells which have acquired the human bladder carcinoma transformation gene should also contain a large array of other co-transfected human DNA fragments, many of which should bear a copy of a highly repeated (300,000 per haploid) sequence which is interspersed throughout the human genome¹⁰. This sequence is not detectable in normal mouse DNA. Such sequences can be detected by the Southern blot procedure¹¹ using a cloned probe of the highly repeated human sequence^{10,12}. Figure 1, lanes c–g, reveals the presence of many segments bearing the human repeated sequence in the DNAs of five different cell lines, each derived from a separate focus induced by human bladder carcinoma DNA. DNAs not derived from *bona fide* transfectants show no trace of these acquired human sequences (lanes a, b). The successfully transformed cells are thus descended from the small subpopulation of cells in the initially transfected culture⁹ which was able to take up and fix large amounts of donor DNA. We conclude that the transformation of these cells was a direct consequence of the acquisition of sequences of a human bladder carcinoma DNA. This or similar hybridization tests may be used as assays to distinguish *bona fide* transfectants from other

Table 1 Transfection of DNAs of glioma and neuroblastoma cell lines

Donor DNA	Tissue type, strain and species	Mode of tumour induction	Efficiency of primary transfection (FFU per 75 µg DNA)	Efficiency of secondary transfection (FFU per 75 µg DNA)	Tumorigenicity of primary transfectant	Retrovirus rescue assay of primary transfectant
G26-20	C57BL/6 mouse glial*	3-methylcholanthrene precipitate was implanted directly into mouse brain	10†	16†, 22†	ND	ND
NIH3T3	NIH/Swiss embryo		0†		1/10‡, 1/7‡	0
B50	BDIX rat, neurone§	Fetal rats were exposed to nitrosoethylurea transplacentally	6, 9, 13, 16	78, 82†, 100†	8/8	0
B104	Same		9, 31	500	6/6, 7/7	0
B103	Same	by injection into pregnant mother on day 15 after conception	93, 122†	12†, 12†	9/9	0
B65	Same		0, 0–1, 0†			
B15	BDIX rat, glial		0†, 0†, 1			

One transfection (fourth and fifth columns) is defined as 75 µg DNA applied to 1.5×10^6 NIH3T3 cells and values represent total number of foci per transfection. Transfection procedures were as described elsewhere¹. Tumorigenicity of transfectant clones was determined by injecting subcutaneously $1-2 \times 10^6$ cells suspended in 0.1 ml PBS into 1–2-day-old newborn mice of NFS or NIH/Swiss outbred strains. Tumour formation was seen after 2 weeks. NIH3T3 cells were injected in an identical manner as a control. Retrovirus rescue assays were designed to induce the release of any defective, transforming viruses which might be present in the transfected cultures as described elsewhere¹. The titre of helper virus released from infected cultures in these rescue assays was routinely 1×10^7 plaque forming units per ml. Infection of Moloney murine sarcoma virus-transformed NIH3T3 non-producer cells with helper virus, performed in parallel as positive control, resulted in release of $>5 \times 10^3$ focus forming units (FFU) of transforming virus per ml. Lines B65 and B15 were used as negative controls for the transfection assay. ND, not determined. IMR32 human neuroblastoma¹³ and NB2A (c1300) mouse neuroblastoma¹⁴ DNAs did not induce a significant number of foci.

* See ref. 15 for origin.

† Not performed in double-blind.

‡ These tumours were seen only at 4 weeks post-inoculation.

§ See ref. 3 for origin.

Table 2 Transfection of DNAs of carcinoma and other tumour lines

Donor DNA	Tissue type, strain and species	Mode of tumour induction	Efficiency of primary transfection (FFU per 75 µg DNA)	Efficiency of secondary transfection (FFU per 75 µg DNA)	Tumorigenicity of primary transfectants	Retrovirus rescue assay of primary transfectant
Lewis lung carcinoma tumour	C57BL/6 mouse lung epithelial*	Spontaneously occurring in animal	7†, 12, 13, 66	95, 105†	9/9	0
Liver of Lewis lung carcinoma-bearing mice	C57BL/6 mouse liver		1-2, 4			
MB48BT-26	C57BL/6 mouse bladder epithelial‡	Isolated bladder epithelial explant was exposed to 7, 12-dimethylbenzanthracene <i>in vitro</i>	48, 56†	10†	8/8	0
NIH3T3	NIH/Swiss mouse embryo fibroblast		0, 0, 0-1		0/9, 0/6	
RBC	NZW rabbit bladder, epithelial§	As above, except benzo-a-pyrene was used	26, 101, 165	52	7/7	0
Rabbit bladder	NZW rabbit bladder		1, 1			
EJ	Human bladder, epithelial	Spontaneously occurring	40†, 66†, 70†	83†, 100†	9/9, 15/15	0
Human epidermal cell strain	Human epidermis		0†, 0†			

DNA transfection procedures were as described elsewhere¹; one transfection is defined as 75 µg DNA for 1.5×10^6 NIH3T3 recipients. Tumorigenicity and retrovirus rescue assay were carried out as described in Table 1 legend. The DNAs of the following tumour cell lines induced few or no foci in a transfection assay: SCC9, SCC12X, SCC4 squamous cell carcinoma cell lines¹⁶; A1663 a human bladder carcinoma line¹⁷; A549, a human lung carcinoma line¹⁷; Detroit 562, a human pharyngeal carcinoma line¹⁸; RT4, a human bladder carcinoma line¹⁹; BC6 a rat bladder carcinoma line (obtained from Dr Howard Green, MIT); and Nettesheim squamous cell carcinoma (obtained from A. D. Little Co.). The DNAs of other tumour cell lines induced few or no foci in a DNA transfection assay: HTC rat hepatoma established by Dr Gordon Tomkins; 7777 rat hepatoma²⁰; human Murawski melanoma and human Wolfe melanoma (obtained from Dr Lloyd Old); A375 human melanoma¹⁷; M3 mouse melanoma²¹ and B16 mouse melanoma²².

*For information on origin of cell line see Dr M. R. Lewis, unpublished data.

† Not performed in double-blind.

‡ For origin of cell line see ref. 6.

§ See ref. 5.

|| Origin of cell line Dr J. Daly, unpublished data.

colonies in cultures which have become transformed by mechanisms other than transfection.

The positive results with the human, rabbit and rat tumour DNAs indicate that tumour transforming genes are able to act across species barriers—they may also act across tissue barriers: for example, transforming sequences of neuroblasts and epithelial cells can transform the recipient fibroblasts. In addition, transforming sequences that are active in these transfections (Tables 1, 2) may be derived from tumours induced by a variety of carcinogens which have been administered in different ways. We conclude that the transfection assay can reveal transformation sequences in tumours of diverse origin and is therefore not limited to analysis of donor tumour DNAs of mouse fibroblast origin. The demonstration of these transmissible genes from various tumours will allow their molecular cloning

and the analysis of the molecular basis of non-viral carcinogenesis.

We thank the following people for providing tumour cell lines: Drs I. Summerhayes, L. B. Chen, J. Rheinwald, H. Green, O. Kehinde, J. Patrick, T. Yean-Kai, I. Wodinsky, G. Todaro,

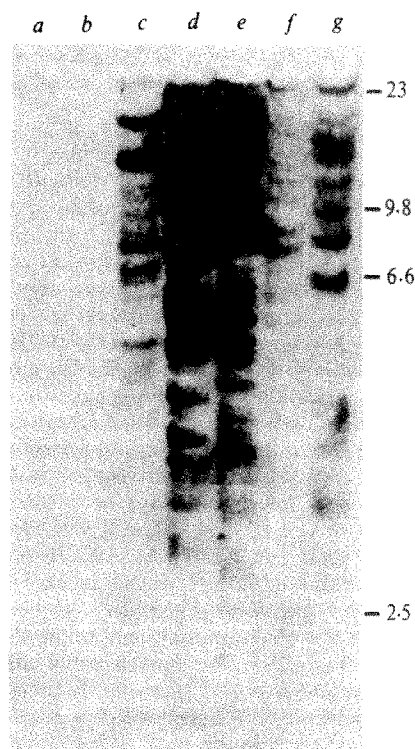


Fig. 1 Detection of human DNA sequences in transformed NIH3T3 cells following transfection with EJ carcinoma DNA. DNA was prepared from five foci; 10 µg of each DNA was digested with *Eco*RI and electrophoresed through a 1% agarose gel in 40 mM Tris (pH 7.9), 50 mM sodium acetate and 1 mM EDTA. The gel was run at 1.2 V cm^{-1} for 16 h. After electrophoresis, the DNA was transferred to nitrocellulose by the method of Southern¹¹, and the resulting blot hybridized with the BLUR8 clone of the human repetitive 'Alu family' DNA¹². The Alu probe was ³²P-labelled by nick translation²³ and 5×10^6 c.p.m. of labelled probe was hybridized to the blot in the presence of 10% dextran sulphate²⁴. Hybridization was carried out for 40 h followed by a 12-h exposure to X-ray film. Whole cellular human DNA used as a positive control gave extensive blackening throughout the entire lane, without discernible bands (not shown here). Approximate size ranges (in kilobases) are indicated to the right of the gel. a, b, DNAs of untransformed NIH3T3 cells, picked and cloned from dishes which also contained transformed foci. c-g, DNAs from different transformed foci of NIH3T3 cells, derived from transfection of EJ donor DNA.

L. Old, J. Fogh, T.-W. Chang and C. Cepko; Drs P. Deininger, C. Schmid and W. Jelinek for the BLUR8 clone; Dr James Murphy for histology of several tumour samples. We also thank A. Dannenberg and H.-W. Hsu for technical assistance and Dr B. Shilo for useful discussions. This work was supported by funds from the Rita Allen Foundation of New York and the NIH, including core grant CA14051 and program project CA26717. C.S. was supported by a training grant to the MIT Department of Biology, L.C.P. was the recipient of an Eleanor Roosevelt International Fellow and M.M. was a fellow of the Damon Runyon-Walter Winchell Foundation.

Received 22 September 1980; accepted 13 January 1981.

- Shih, C., Shilo, B., Goldfarb, M. P., Dannenberg, A. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5714 (1979).
- Graham, F. L. & van der Eb, A. J. *J. Virol.* **52**, 456 (1973).
- Schubert, D. *et al. Nature* **249**, 224 (1974).
- Dawson, C. T., Sundarraj, N. & Pfeiffer, S. E. *J. biol. Chem.* **252**, 2777 (1977).
- Summerhayes, I. C., Sun, T. T. & Chen, L. B. *J. Cell Biol.* (in the press).
- Summerhayes, I. C. & Franks, L. M. *J. natn. Cancer Inst.* **62**, 1017 (1979).
- Marshall, C. J., Franks, L. M. & Carbonell, A. W. *J. natn. Cancer Inst.* **58**, 1743 (1977).
- Sugiura, K. & Stock, C. C. *Cancer Res.* **15**, 38 (1955).
- Wigler, M. *et al. Cell* **16**, 777 (1979).
- Rubin, C. M., Houck, C. M., Deininger, P. L., Friedmann, T. & Schmid, C. W. *Nature* **284**, 372 (1980).
- Southern, E. M. *J. molec. Biol.* **98**, 503 (1975).
- Jelinek, W. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1398 (1980).
- Tumilowicz, J. J., Nichols, W. W., Cholon, J. J. & Green, A. E. *Cancer Res.* **30**, 2110 (1970).
- Klebe, R. J. & Ruddle, F. H. *J. Cell Biol.* **43**, 69 (1969).
- Sugiura, K. in *Methods of Animal Experimentation* Vol. 2 (ed. Gay, W. L.) 171-222. (Academic, New York, 1969).
- Rheinwald, J. G. & Beckett, M. A. *Cell* **22**, 629 (1980).
- Giard, D. J. *et al. J. natn. Cancer Inst.* **51**, 1417 (1973).
- Peterson, W. D. Jr, Stulberg, C. S. & Simpson, W. F. *Proc. Soc. exp. Biol. Med.* **136**, 1187 (1971).
- Rigby, C. C. & Franks, L. M. *Br. J. Cancer* **24**, 746 (1970).
- McMahon, J. B., Kelleher, P. C. & Smith, C. J. *Biochem. biophys. Res. Commun.* **76**, 1144 (1977).
- Yasumura, Y., Tashjian, A. H. & Sato, G. H. *Science* **154**, 1186 (1966).
- Handbook on Genetically Standardized Jax Mice* (Jackson Memorial Laboratory, Bar Harbor, 1962).
- Rigby, P. W. J., Dickmann, M., Rhodes, C. & Berg, P. J. *molec. Biol.* **113**, 237 (1977).
- Wahl, G., Stern, M. & Stark, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683 (1979).

Identification of mutations affecting replication control of plasmid Clo DF13

Antoine R. Stuitje, Cornelis E. Spelt, Eduard Veltkamp & H. John J. Nijkamp

Department of Genetics, Biological Laboratory, Vrije Universiteit, De Boelelaan 1087, 1007 MC Amsterdam, The Netherlands

The bacteriocinogenic plasmid Clo DF13, originally isolated from *Enterobacter cloacae*¹, is stably maintained in *Escherichia coli*² to the extent of about 10 copies per cell³. Its replication⁴ resembles that of many other small, multicopy plasmids⁵⁻⁸; plasmid-encoded protein is not required⁹⁻¹² but plasmid-specific genetic information is involved in regulation of replication as both conditional¹³ and nonconditional³ copy-number mutants of Clo DF13, and transcomplementable^{14,15} copy-number mutants of plasmid Col E1 have been described. The sequences essential for replication of Col E1 (refs 16, 17) and Clo DF13 (refs 18, 19) have been identified within a region surrounding the replication origin. Initiation of Col E1 replication is preceded by transcription of the origin region, providing the RNA primer at the origin²⁰. However, transcription in the opposite direction results in a small transcript of ~100 nucleotides (RNA-100) for both Col E1 (refs 21, 22) and Clo DF13 (ref. 23). Data suggest that Col E1 RNA-100 acts as a negative control element for the initiation of replication^{20,24}. We show here that single base transitions in the RNA-100 cistron of Clo DF13 can result in a nonconditional increase in plasmid copy-number. Also, sequence analysis has revealed that a specific base transition in a DNA region, apparently involved in both termination and initiation of transcription towards the replication origin, results in a thermosensitive plasmid copy-number.

Both nonconditional³ and conditional¹³ replication-control mutants of plasmid Clo DF13 (8.6 kilobases) have been isolated after chemical mutagenesis, and characterized^{13,25}. As reported elsewhere¹⁸, the Clo DF13-*cop3* mutation (Table 1) was mapped by insertion of transposons and successive deletion of specific restriction fragments. Similar studies on the Clo DF13-*cop1* (Ts) and the Clo DF13-*cop2* mutants (Table 1) show that these mutations are also located within the region (1.8-15%) surrounding the replication origin. These copy mutations were localized more precisely within this region by *in vitro* recombination experiments (see Fig. 1 legend). Subsequently, specific regions of the different copy mutants were sequenced as indicated in Table 1. By comparing these sequences, the complete nucleotide sequence between the *Hae*II site at 1.8% and the *Bam*HI site at 15% was deduced (Fig. 2). Furthermore these studies show that these different *cop* mutations are, in fact, point mutations, and all represent base transitions (Fig. 2).

Recently, Itoh and Tomizawa²⁰ demonstrated that the initiation of DNA replication of plasmid Col E1 *in vitro* is preceded by transcription of the origin region, towards the origin of replication, generating a precursor primer RNA molecule (Fig. 1). Furthermore, they showed that when transcription passes the origin, the RNA transcript is able to form a stable hybrid with this region. In the next step RNase H specifically cleaves the hybridized RNA, producing the RNA-DNA hybrid with a free 3' OH end for the initiation of DNA replication by DNA polymerase I. The 5' end of the 'pre-primer' RNA has been sequenced and the corresponding region within the sequence of plasmid Clo DF13 is indicated in Fig. 2.

Figures 2 and 3 show that the Clo DF13 *cop1* (Ts) mutation involves a base transition at position 720. The region surrounding this mutation has several interesting features: first, the nucleotide sequence of this region (see Fig. 4b) resembles that of regions known to be involved in termination of transcription²⁶. A similar role for the region in Clo DF13 is supported by data²³ which shows that transcription from the cloacin promoter results in the synthesis of two RNA molecules of 2.2 and 2.4 kilobases respectively. The terminator for the transcription resulting in the synthesis of the 2.4-kilobase RNA was located at 9.3%, and therefore maps within the sequence mentioned above (T2, Fig. 1). Second, the nucleotide sequence of the region surrounding the *cop1* (Ts) mutation (position 696-735) is completely conserved when compared with the corresponding sequence of plasmid Col E1 (ref. 16). The position of the initiation site of transcription of the Col E1 pre-primer RNA, as determined by Itoh and Tomizawa²⁰, implicates the Col E1 region mentioned above in promotion of the synthesis of this RNA. Therefore, the corresponding region of plasmid Clo DF13 which surrounds the *cop1* (Ts) mutation might also act as promoter for RNA synthesis (P3, Fig. 1). In view of these data, the *cop1* (Ts) mutation could be a promoter-type or terminator-type mutation. Alternatively, it could be a mutation at the site of action of a control element. Further experiments are in progress to discriminate between these possibilities and determine whether the effect of the *cop1* (Ts) mutation, a temperature-inducible increase of the plasmid copy-number, is caused by an increased synthesis of pre-primer RNA at the elevated temperature.

The nonconditional *cop2* and *cop3* mutations are located within a Clo DF13 region which is transcribed *in vivo* and *in vitro* resulting in a RNA molecule of ~100 nucleotides^{19,23} (RNA-100, see Figs 2, 3). The corresponding region of plasmid Col E1 is also transcribed *in vitro*^{21,22}, producing a similar molecule. The nucleotide sequence of the latter RNA-100 species has been determined and is indicated in Fig. 4a. Perusal of the Clo DF13 and Col E1 sequences encoding this RNA-100 species strongly suggest that these RNA molecules are not translated. Furthermore, it was observed that although the primary structures of the Clo DF13 and Col E1 RNA-100 species differ largely, the secondary structure seems to be highly conserved (Fig. 4a). The position of the *cop2* and *cop3* mutations within the sequence (Fig. 2) makes it unlikely that these mutations directly affect the synthesis of either the RNA-100 or the pre-primer RNA.

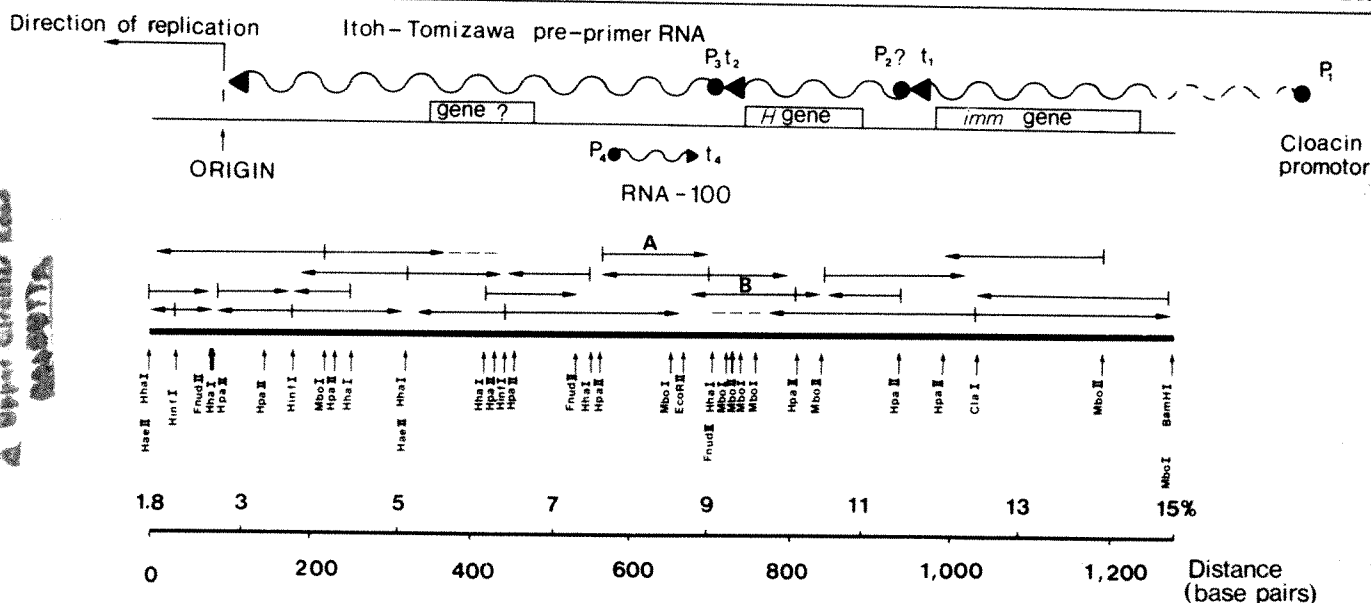


Fig. 1 Detailed physical/genetic map of the region between the *HaeII* site and the *BamHI* site at respectively 1.8 and 15% on the Clo DF13 map¹⁹. The positions of the various restriction endonuclease cleavage sites were determined as reported previously¹⁹ and verified by sequence analysis. The exact position of the immunity gene (*imm*) could be determined by DNA sequence analysis of this region, as part of the amino acid sequence of the Clo DF13 immunity protein has been elucidated²⁷. The region of the Clo DF13 plasmid genome encoding the H protein was previously described²³. The nucleotide sequence probably encoding this polypeptide of about 50 amino acids is indicated in Fig. 2. The region possibly encoding an arginine-rich, basic polypeptide of about 45 amino acids, for both Col E1 and Clo DF13 (ref. 19), is indicated as 'gene?'. As reported²³, transcription initiated at the cloacin promoter (P1) terminates at t1 or t2 resulting in RNA molecules of ~2,200 and 2,400 nucleotides respectively. By comparing the Clo DF13 and Col E1 origin sequences, the position of the Clo DF13 RNA-100 cistron was deduced¹⁹. As indicated, the origin of replication maps within the 1.8–5% *HaeII* restriction fragment²⁸. By partial cleavage of a Clo DF13-*cop3* derivative (pEV26)¹⁸ with *HaeII* endonuclease in the presence of S₁ nuclease, followed by ligation with T₄ ligase and transformation of *E. coli*, we have isolated a plasmid which has an altered *HaeII* (1.8–5%) 'origin' fragment and still contains the *cop3* mutation (A.R.S. *et al.*, unpublished results). Recombinant plasmids containing both wild-type (wt) and *cop3 HaeII* fragments were obtained by transforming cells to ampicillin resistance with a ligated mixture of wild-type and *cop3 HaeII* fragments. The composition of the recombinant plasmids was easily determined from the *HaeII* digestion patterns. The Cop phenotype was deduced from the minimal inhibitory concentration of ampicillin of cells harbouring these plasmids. By these types of study the Clo DF13-*cop3* and -*cop2* mutations were localized more precisely within the 1.8–15% origin region (see Table 1).

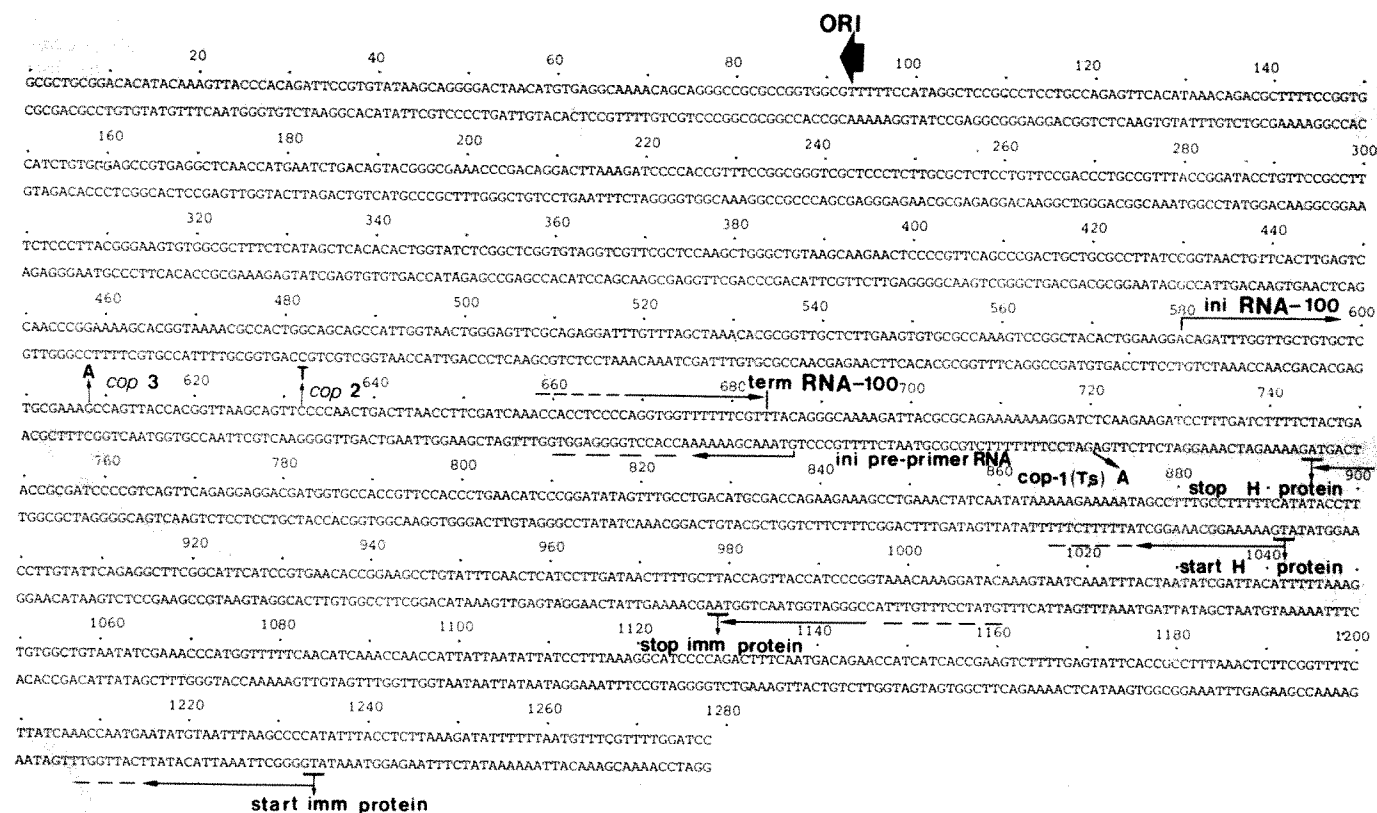


Fig. 2 The nucleotide sequence of the Clo DF13 region between the *Hae*II site at 1.8% and the *Bam*HI site at 15% on the Clo DF13 physical map (see also Fig. 1). The sequencing procedures were as described previously¹⁹. The sequence strategy is indicated in Fig. 1 legend; the arrows in Fig. 1 show the part of the restriction fragments sequenced. The direction of the arrows indicates the strands which were sequenced, aligned in 5' to 3' direction (tail to head). Start and stop indicate respectively the start and stop codons for proteins; 'ini' and 'term' represent the 5' and 3' ends of RNAs respectively.

Table 1 Properties of Clo DF13 and mutant plasmids

Plasmid	Derivation	Relevant properties	Copies per cell	MW (kilobases)	Clo DF13 region present (%)	Map* position of the <i>cop</i> mutation (%)	Clo DF13 region of plasmid sequence (%)	Ref. source
Clo DF13	Naturally occurring in <i>E. cloacae</i>	Cloacin DF13	10	8.6	0–100	—	—	1
Clo DF13- <i>cop3</i>	Clo DF13	Cloacin DF13, nonconditional copy mutant	70	8.6	0–100	5–11.5	—	21
Clo DF13- <i>cop2</i>	pEV5	Nonconditional copy mutant, Ap	200	2.9	1.8–15	5–15	5–15	Stuitje (unpublished)
Clo DF13- <i>cop1</i> (Ts)	pJN03	Cloacin DF13, ApTs-copy mutant Ts-kill	Temperature dependence	13.1	0–100	1.8–9.5		14
pEV5	Clo DF13::Tn901 insertion mutant	Ap	10	2.9	1.8–15		5–15	Stuitje (unpublished)
pEV22	Clo DF13- <i>cop3</i> ::Tn901 insertion mutant	Ap	200	2.8	1.8–11.5		1.8–11.5	4
pVU206	Clo DF13- <i>cop1</i> (Ts)	Ap Ts-copy mutant Ts-kill	Temperature dependence	3.3	1.8–15 27–29		1.8–15	Stuitje (unpublished)

Ts, temperature sensitive.

* The different *cop* mutations were localized more precisely within the 1.8–15% origin region by recombining the *Hae*II 'ori' fragments of the different mutant plasmids, as described in Fig. 1 legend.

Although our data on the *cop2* and *cop3* mutations do not exclude a possible role of the RNA-100 as a positive control element, they are consistent with previous findings that favour the RNA-100 as a negative control element for the initiation of DNA replication^{20,24}. Furthermore, results²⁵ suggest that the RNA-100 is involved in plasmid incompatibility, a plasmid function that seems to be intimately connected with replication control. In fact, it has been reported²⁰ that the presence of the RNA-100 species of plasmids belonging to the Col E1 incompatibility group inhibits the formation of the primer *in vitro*. The sequence data described in this letter indicate that there are at least two distinct regions that are important in the control of Clo DF13 replication. The first encodes a RNA-100 molecule whose synthesis is initiated ~490 nucleotides upstream from the replication origin, whereas the second, which is located about 620 nucleotides upstream from the origin,

seems to be involved in both termination and initiation of transcription towards the origin. The identification of the different mutations inevitably evokes questions on the mechanism of replication control of small plasmids such as Col E1, for example: (1) is the frequency of initiation of DNA replication controlled at the level of pre-primer RNA synthesis; (2) is the synthesis of pre-primer RNA regulated at P3/t2 (see Fig. 1); (3) in which way does the RNA-100 species control plasmid replication and are, in addition to this RNA, other plasmid-encoded gene products involved? Studies on additional copy-control mutants, including complementation experiments between the mutants and studies on the effects of alterations generated *in vitro* by substitution or deletion of specific sequences at different regulatory sites, may elucidate the actual mechanism of plasmid DNA replication control.

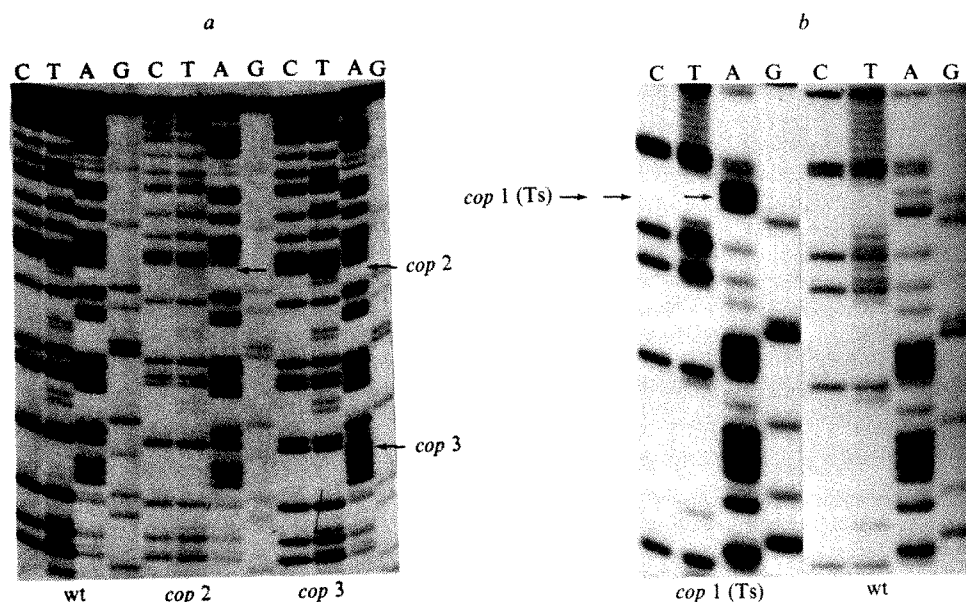


Fig. 3 *a*, Result of the Gilbert-Maxam sequencing procedure applied to fragment A (labelled at the 5' end of the *Hpa*II site as described previously¹⁹; see Fig. 1) derived from the wild-type (wt) Clo DF13 and the *cop2* and *cop3* mutants. Note the depression in the wild-type sequence near the position of the *cop3* mutation. The actual wild-type sequence 5'-A-A-A-G-C-C-A-3', instead of 5'-A-A-A-G-C-A-3', was deduced by sequence analysis of the complementary strand. *b*, Result of the Maxam-Gilbert sequencing procedure applied to fragment B (5' end-labelled at the *Hpa*II site; see Fig. 1) of the wild-type Clo DF13 plasmid and the Clo DF13-*cop1* (Ts) mutant.

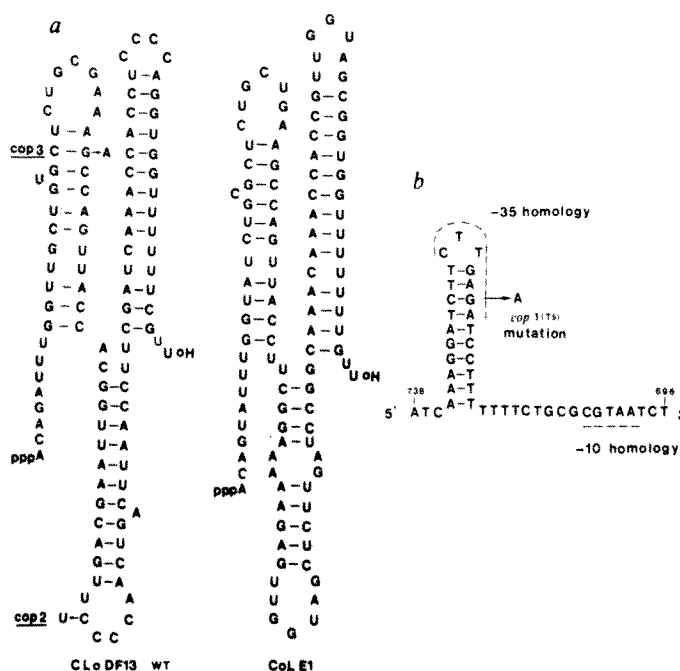


Fig. 4 *a*, Comparison of the secondary structures of the RNA-100 species encoded by the plasmids Col E1 and Clo DF13. The resident *cop2* (C → U) and *cop3* (G → A) mutations are indicated. The most stable secondary structure was estimated by the procedure described by Tinoco *et al.*²⁹ and the ΔG (25°C) values estimated for the wild-type (wt) Clo DF13, *cop2*, *cop3* and Col E1: 100-nucleotide RNAs are respectively 50.4, -50.2, -41.2 and -45.6 kcal. *b*, Secondary structure of the DNA region surrounding the *cop1* (Ts) mutation (G → A transition). This region is presumed to be involved in both termination and initiation of transcription towards the origin (transcription proceeds from left to right). The sequences which are supposed to be involved in recognition and binding of RNA polymerase are indicated by -35 and -10 homology respectively²⁶.

This study was supported in part by The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands Organization for Advancement of Pure Science.

Received 10 October 1980; accepted 2 January 1981.

1. Tieze, G. A., Stouthamer, A. H., Jansz, H. S., Zandberg, J. & van Bruggen, E. F. J. *Molec. gen. Genet.* **106**, 48–65 (1969).
2. Kool, A. J., Pranger, M. & Nijkamp, H. J. J. *Molec. gen. Genet.* **115**, 314–323 (1972).
3. Kool, A. J. & Nijkamp, H. J. J. *Bact.* **120**, 569–578 (1974).
4. Veltkamp, E. & Stuitje, A. R. *Plasmid* (in the press).
5. Staudenbauer, W. L. *Curr. Topics Microbiol. Immun.* **83**, 93–157 (1978).
6. Bolivar, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 5265–5269 (1977).
7. Conrad, S. E., Wold, M. & Campbell, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 736–740 (1979).
8. Chang, A. C. Y. & Cohen, S. N. J. *Bact.* **134**, 1141–1156 (1978).
9. Kahn, M. & Helinsky, D. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2200–2204 (1978).
10. Donoghue, D. J. & Sharp, P. A. *J. Bact.* **133**, 1287–1294 (1978).
11. Tomizawa, J. *Nature* **257**, 253–254 (1975).
12. Staudenbauer, W. L. *Molec. gen. Genet.* **145**, 273–280 (1976).
13. Andreoli, P. M., Brandsma, R. A., Veltkamp, E. & Nijkamp, H. J. J. *Bact.* **135**, 612–621 (1978).
14. Twigg, A. J. & Sherratt, D. *Nature* **283**, 216–218 (1980).
15. Shepard, H. M., Gelfand, D. H. & Polisky, H. *Cell* **18**, 267–275 (1979).
16. Oka, A. *et al. Molec. gen. Genet.* **172**, 151–159 (1979).
17. Backman, K., Betlach, M., Boyer, H. W. & Yanowski, S. *Cold Spring Harb. Symp. quant. Biol.* **43**, 69–76 (1979).
18. Stuitje, A. R., Veltkamp, E., van den Elzen, P. J. M. & Nijkamp, H. J. J. *Nucleic Acids Res.* **5**, 1801–1820 (1978).
19. Stuitje, A. R., Veltkamp, E., Maat, J. & Heijneker, H. L. *Nucleic Acids Res.* **8**, 1459–1473 (1980).
20. Itoh, T. & Tomizawa, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2450–2454 (1980).
21. Patient, R. K. *Nucleic Acids Res.* **6**, 2647–2666 (1979).
22. Morita, M. & Oka, A. *Eur. J. Biochem.* **97**, 435–443 (1979).
23. Van den Elzen, P. J. M., Konings, R. N. H., Veltkamp, E. & Nijkamp, H. J. J. *Bact.* **144**, 579–591 (1980).
24. Conrad, S. E. & Campbell, J. L. *Cell* **18**, 61–71 (1979).
25. Hashimoto-Gotoh, T. & Inselberg, J. *J. Bact.* **139**, 608–619 (1979).
26. Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319–353 (1979).
27. Van den Elzen, P. J. M. *et al. Nucleic Acids Res.* (in the press).
28. Stuitje, A. R., Veltkamp, E., Weyers, P. J. & Nijkamp, H. J. J. *Nucleic Acids Res.* **6**, 71–80 (1979).
29. Tinoco, I. *et al. Nature new Biol.* **246**, 44, 40–41 (1973).

Differential usage of iso-accepting tRNA^{Ser} species in silk glands of *Bombyx mori*

Danièle Hentzen, Anna Chevallier & Jean-Pierre Garel

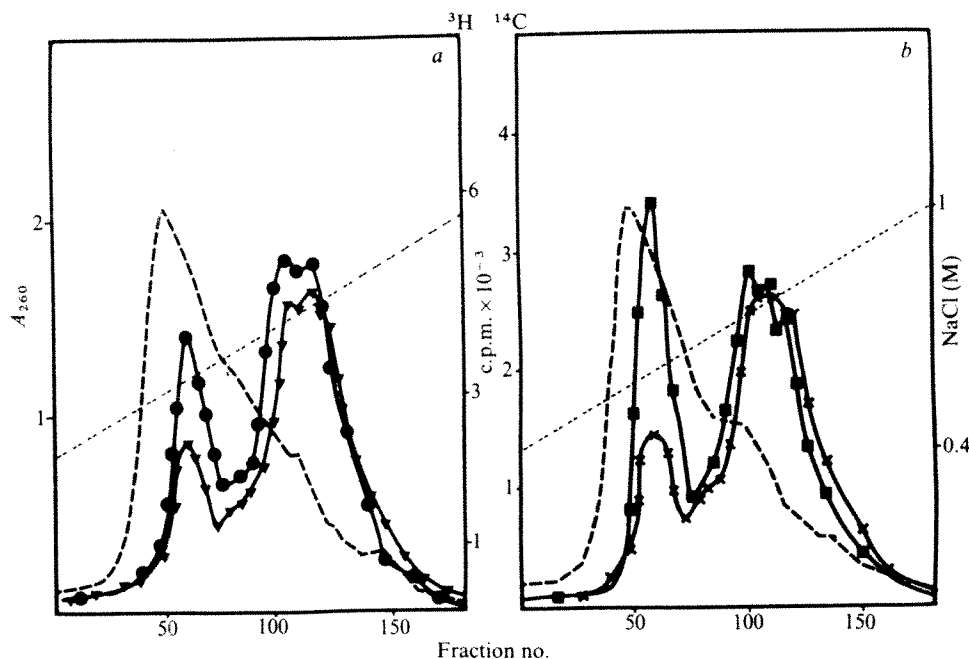
Department of General and Applied Biology, Université Claude Bernard, Lyon I, 43, Bd du 11 Novembre 1918, F-69622 Villeurbanne Cedex, France

The rapid development of the silk glands of *Bombyx mori* during the last larval instar shows two phases. During the first 4 days, in both the middle and posterior parts of the silk glands, the ribosomal machinery is assembled and the synthesis of house-keeping proteins starts. During the second phase (the last 4 days), the middle part of the gland synthesizes ~45 mg of the silk protein sericin (31% serine) and the posterior part of the gland synthesizes ~130 mg of the silk protein fibroin (46% glycine, 29% alanine and 12% serine). Silk fibroin and sericin are detectable by the second day and represent 80 and 50% respectively of the total proteins produced at day 8 (refs 1–4). It is known that the tRNA population of the posterior part of the gland is quantitatively adapted to fibroin codon frequency during this period^{5–9} but little is known about the situation in the middle part except for the observation that it contains more tRNA^{Ser} than does the posterior part. We show here that the two parts contain, and presumably use, different iso-accepting species of tRNA^{Ser}, the middle part using tRNA₁^{Ser}, which recognizes AGU and AGC codons, and the posterior part using tRNA₂^{Ser} which recognizes UCA. We also suggest that this differential adaptation of the tRNA^{Ser} species is under transcriptional control as the two species are accumulated at different rates, but degraded at the same rate.

tRNA^{Ser} species were fractionated and quantitated using benzoylated DEAE (BD)-cellulose chromatography¹⁰. Figure 1 shows chromatographic profiles of acylated Ser-tRNA from middle and posterior silk glands of *B. mori* at two different stages of development: the growth phase (day 2) and at the end of silk secretion (day 8). A striking difference occurs in the distribution of the subsets of tRNA^{Ser} species: in the fibroin-producing part, the tRNA₂^{Ser} group is predominant (82% during the secretion phase) but in the sericin-producing compartment, the tRNA₂^{Ser}/tRNA₁^{Ser} ratio changes from 3.1 at the start of the growth phase to 1.8 during the secretion phase. The tRNA₁^{Ser} species then represents 36% of the iso-tRNA^{Ser} family. From base composition, codon recognition and partial sequencing, we know that the tRNA₁^{Ser} species is capable of translating AGU and AGC codons, whereas the tRNA₂^{Ser} group recognizes UCA codons^{6,11,12}.

Refinements were made using two-dimensional polyacrylamide gel electrophoresis^{13,14}. As shown in Fig. 2, five spots can be assigned to iso-tRNA^{Ser} species. The tRNA₂^{Ser} subset fractionates according to the diagonal into three species: two tRNA_{2a}^{Ser} subspecies with the lowest mobility and one tRNA_{2b}^{Ser} with the highest mobility (Fig. 2*a*); tRNA₁^{Ser} migrates with an intermediate mobility. It seems to be heterogeneous and resolves into two distinct spots (Fig. 2*b–d*). Cross-identifications using purified tRNA^{Ser} species collected from a specifically retarded elution on BD-cellulose chromatography, done in the absence of Mg²⁺ ions¹⁵, yield the final pattern shown in Fig. 2*c, d*. The differential accumulation of iso-tRNA^{Ser} species is made clear on the electrophoretic maps in Fig. 3. It can be seen that the tRNA₁^{Ser} species is more abundant in the middle silk gland (MSG) than in the posterior silk gland (PSG) during the silk secretion phase. Polyacrylamide electrophoresis corroborates the BD-cellulose chromatographic profiles. An unexpected result was the unusual composition of the tRNA₂^{Ser} subset. tRNA_{2b}^{Ser} carrying the IGA anticodon is usually the major

Fig. 1 Chromatography on BD-cellulose of Ser-tRNA from posterior (PSG) and middle (MSG) silk glands of *B. mori* silkworm. *a*, Elution profiles of tRNA extracted from the middle part on the second day (growth phase) of the fifth larval instar and acylated with ^3H -serine ($\bullet$), and tRNA from the posterior part acylated with ^{14}C -serine ($\blacktriangledown$). The co-chromatography was achieved on a 40×2.2 cm BD-cellulose column eluted with a linear gradient of 0.2–1.0 M NaCl in 2 l of buffer: 10 mM sodium acetate, pH 4.7, 10 mM MgCl_2 at 25°C . 10-ml fractions were treated with 0.1 volume of 55% trichloroacetic acid in the presence of 0.5 mg carrier tRNA. The precipitate was filtered on to a glass fibre disk and radioactivity measured in a liquid scintillation photometer. This experiment was repeated twice after reversing the labelling. Absorbance was measured at 260 nm (---). *b*, Similar experiments with tRNA extracted on the eighth day of the fifth instar: ($\times$) PSG tRNA acylated with ^3H -serine; ($\blacksquare$) MSG tRNA acylated with ^{14}C -serine.



species among the eukaryotic tRNA^{Ser} family able to decode UCU, UCC and, with less efficiency, UCA; a minor tRNA^{Ser}_{2c} species is able to recognize the rare UCG codon^{15,16}. The tRNA^{Ser}_{2a} species, which accumulates at a higher level in both silk gland parts than in other non-silk producing tissues (see the carcass in Fig. 3c), might be able to decode specifically the UCA codon—the major serine codon found in fibroin mRNA¹⁷.

A side conclusion can be drawn from these results, based on differential accumulation of specific iso-tRNA^{Ser} species related to the developmental stage of a tissue. The quantitative adaptation of tRNA species to codon frequency demonstrated for a specific and typical mRNA—fibroin or globin¹⁶—leads us to

suggest that the codon usage for serine in the sericin message is somewhat different from that found in fibroin mRNA. The high level of tRNA^{Ser}₁ in the MSG can be easily explained if we state that sericin mRNA carries AGU and/or AGC codons in addition to the UCA codon.

The specific requirement of the iso-tRNA^{Ser} family in decoding different mRNAs in two distinct silk gland parts remains to be understood. We have studied tRNA synthesis using *in vivo* short pulse-labellings of tRNA throughout the last larval instar and then analysed the tRNA species on electrophoretic maps. Synthesis rates of tRNA are not uniform: rapid changes occur, which are specific for each tRNA species and for each silk gland

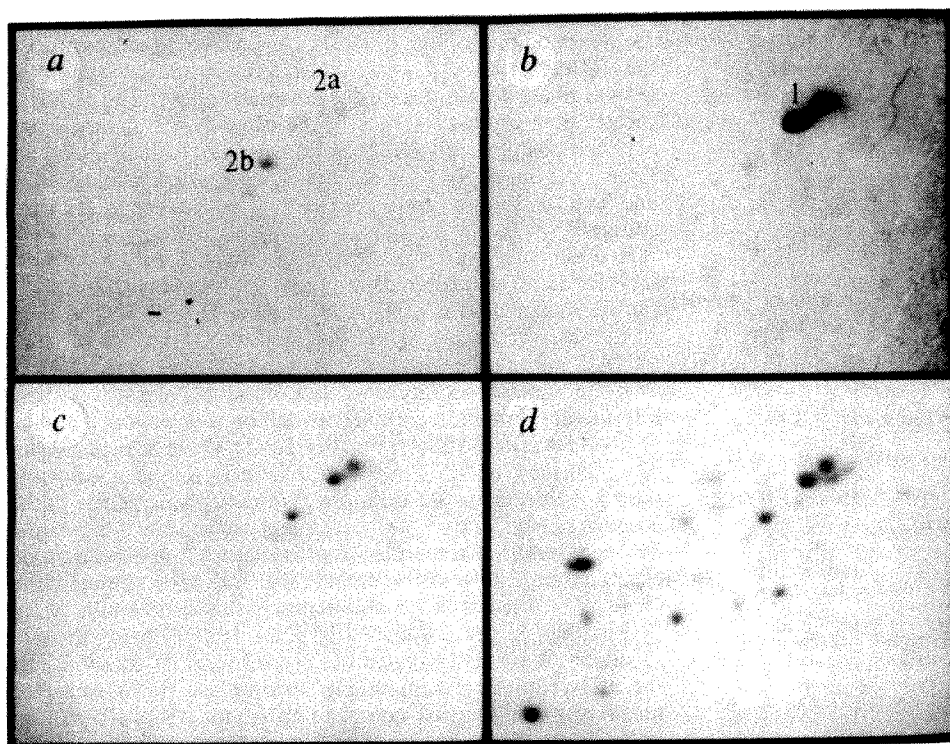


Fig. 2 Identification of tRNA^{Ser} species from the MSG (eighth day from the fifth instar) fractionated on a BD-cellulose column and analysed by two-dimensional polyacrylamide gel electrophoresis. The first dimension was carried out on 10% acrylamide–bisacrylamide (19:1), 7 M urea, Tris-borate buffer 0.089 M, pH 8.3, and run from right to left for 25 h at 12 V cm^{-1} in the cold. The second dimension was done on 20% acrylamide, 4 M urea in the same buffer and run from top to bottom for 48 h at 13 V cm^{-1} . Gels were stained in 0.2% (w/v) methylene blue solution in 2% sodium acetate buffer, pH 5.0, and destained in running water^{13,14}. *a*, tRNA^{Ser}₂ species from the second peak of the BD-cellulose column (fractions 90–150 from Fig. 1b). *b*, tRNA^{Ser}₁ species from the first peak (fractions 45–80). *c*, tRNA^{Ser}₁ and tRNA^{Ser}₂ species previously purified from a BD-cellulose column carried out in the absence of Mg^{2+} . *d*, tRNA^{Ser}₁ and tRNA^{Ser}₂ co-electrophoresed with 0.5 A₂₆₀ of total PSG tRNA.

part. It has been shown, for example, that $\text{tRNA}_{1\text{Gly}}$, $\text{tRNA}_{2\text{Gly}}$ and $\text{tRNA}_{2\text{Ala}}$ are the most highly labelled species at all times in the posterior part. Their absolute synthesis rates are at a maximum on days 3 and 4, that is, at the end of the growth phase and the beginning of fibroin secretion. In the middle part however, tRNA_{Asp} , $\text{tRNA}_{1\text{Gly}}$, $\text{tRNA}_{1\text{Ser}}$ and $\text{tRNA}_{2\text{aSer}}$ species have high absolute synthesis rates with a maximum shifted in time, whereas other tRNA species are synthesized at lower rates.

A comparison of biosynthetic rates for the iso- tRNA_{Ser} family is of particular interest. Figure 4 shows that during the fifth instar changes in the relative synthesis rates for $\text{tRNA}_{1\text{Ser}}$ and $\text{tRNA}_{2\text{aSer}}$ occur in the two silk gland parts—these changes differ between the two parts. In the PSG, there is an increase in $\text{tRNA}_{2\text{aSer}}$ synthesis with respect to $\text{tRNA}_{1\text{Ser}}$ species at the start of fibroin production. However, in the MSG, $\text{tRNA}_{1\text{Ser}}$ species is preferentially synthesized although the changes are more progressive. We know that these different patterns are essentially due to the highest synthesis rate of $\text{tRNA}_{2\text{aSer}}$ in the posterior, and $\text{tRNA}_{1\text{Ser}}$ in the middle part. These individual changes explain

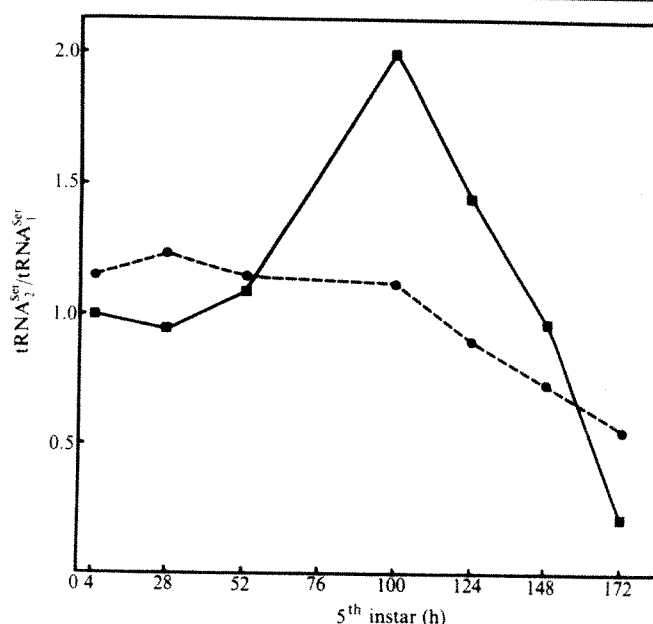


Fig. 4 Differential synthesis rates of $\text{tRNA}_{1\text{Ser}}$ and $\text{tRNA}_{2\text{aSer}}$ species in the middle and posterior silk glands of *B. mori* silkworm during the fifth larval instar. Silkworms from days 1 to 8 of the last instar were each injected with 0.5 mCi ^3H -uridine and dissected after 4 h labelling. tRNA from posterior (■) and middle (●) silk glands were subjected to two-dimensional polyacrylamide gel electrophoresis and each tRNA_{Ser} spot quantitated. The ordinate indicates time (h) after the first meal following the last moult.

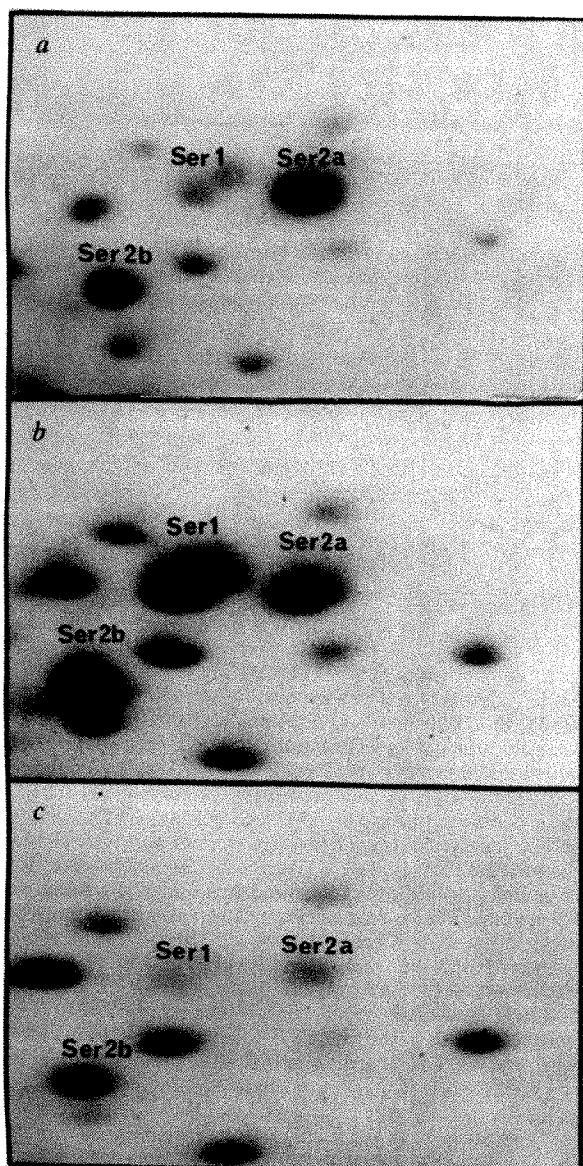


Fig. 3 Differential accumulation of tRNA_{Ser} species on polyacrylamide gel maps from PSG and MSG of *B. mori* silkworm at the end of the fifth larval instar. Two-dimensional electrophoresis was carried out according to Fig. 2 legend. The figures show only the tRNA_{Ser} area corresponding to the upper right part of a complete tRNA electrophoretic map. a, PSG; b, MSG; c, carcass.

the differential accumulations evident in Fig. 1 and shown more precisely in Fig. 3. They seem to be related to the different mRNA populations being decoded.

Of the three possible levels of regulation—preferential degradation of mature tRNA species, selective processing of precursor tRNA, and differential transcription of tRNA genes—we concentrated on the last possibility. The preferential degradation of non-used mature tRNA species can be ruled out, because we have found similar degradation rates for 13 tRNA species, including tRNA_{Ser} species¹³. A selective processing of precursor tRNA, whose primary transcripts would match tRNA gene distribution, cannot be excluded, although preliminary results from silk glands incubated on day 4 for a short time with ^{32}P indicate an enrichment of pre- tRNA_{Ser} in the middle part compared with the posterior part¹³.

Although there is no direct evidence for a transcriptional control of tRNA gene activity according to the needs of the polysomal compartment, the differential appearance and accumulation observed for each iso- tRNA_{Ser} species favour such a mechanism, the most suitable to account for tRNA adaptation to the mRNA codons being translated.

We thank Dr G. Chavancy, A. Fournier and Ms J. Ramsey for critical reading of the manuscript.

Received 14 July 1980; accepted 9 January 1981.

- Daillie, J. C. *hebd. Séanc. Acad. Sci., Paris* **261**, 4872–4875 (1965).
- Shigematsu, H. & Takeshita, H. *J. Insect Physiol.* **14**, 1013–1024 (1968).
- Tashiro, Y., Morimoto, T., Matsuura, S. & Nagata, S. *J. Cell Biol.* **38**, 574–588 (1968).
- Fournier, A. *Biochimie* **61**, 283–320 (1979).
- Daillie, J. *Biol. Cell* **29**, 1–6 (1977).
- Garel, J. P., Hentzen, D. & Daillie, J. *FEBS Lett.* **39**, 359–363 (1974).
- Chen, G. S. & Siddiqui, M. A. Q. *Archs Biochem. Biophys.* **161**, 109–117 (1974).
- Majima, R., Kawakami, M. & Shimura, K. *J. Biochem., Tokyo* **78**, 391–400 (1975).
- Garel, J. P. *Nature* **260**, 805–806 (1976).
- Gillam, I. *et al. Biochemistry* **6**, 3043–3056 (1967).
- Garel, J. P., Hentzen, D., Schlegel, M. & Dirheimer, G. *Biochimie* **58**, 1089–1100 (1976).
- Hentzen, D., Garel, J. P. & Keith, G. *Biochem. biophys. Res. Commun.* **71**, 241–248 (1976).
- Chevallier, A. & Garel, J. P. *Biochimie* **61**, 245–262 (1979).
- Garel, J. P., Garber, R. & Siddiqui, M. A. Q. *Biochemistry* **16**, 3618–3624 (1977).
- Rogg, H., Müller, P. & Staehelin, M. *Eur. J. Biochem.* **53**, 115–127 (1975).
- Chavancy, G., Chevallier, A., Fournier, A. & Garel, J. P. *Biochimie* **61**, 71–78 (1979).
- Suzuki, Y. & Brown, D. D. *J. molec. Biol.* **63**, 409–429 (1972).

Human erythrocyte membranes exhibit a cooperative calmodulin-dependent Ca^{2+} -ATPase of high calcium sensitivity

Peter Downes & Robert H. Michell

Department of Biochemistry, University of Birmingham,
PO Box 363, Birmingham B15 2TT, UK

It is thought that the ionized Ca^{2+} concentration in the cytosol of healthy erythrocytes is in the range 0.01–0.1 μM (ref. 1) and that this low concentration is maintained by an ATP-driven calmodulin-dependent Ca^{2+} pump in the plasma membrane^{2,3}. The Ca^{2+} -stimulated ATPase which is the enzymatic expression of this pump varies in its calcium sensitivity between different preparations of erythrocyte ghosts, with activation generally occurring in a concentration range between ~1 and 10 μM Ca^{2+} (refs 2–5). This is a higher range of Ca^{2+} concentrations than might be anticipated for activation of a pump that sustains intracellular concentrations of Ca^{2+} below 0.1 μM , and recent reports have suggested activation in some membrane preparations at Ca^{2+} concentrations in the range 0.1–1.0 μM (refs 6, 7). We report here a simple method for preparing human erythrocyte membranes in 2.5 mM HEPES/1 mM EGTA at pH 7.0 (see Fig. 1 legend) in which the activation of the Ca^{2+} -ATPase by Ca^{2+} and intracellular concentrations of calmodulin is highly cooperative and is complete by ~1 μM Ca^{2+} . Unlike other available erythrocyte membrane preparations, the pattern of activation by Ca^{2+} and calmodulin is not complicated by partial resealing of the ghosts during and after isolation. We suggest that this cooperative activation of the Ca^{2+} pump may explain how healthy erythrocytes maintain their normal cytosol Ca^{2+} concentration at a threshold value at or below ~0.1 μM . We also note that several other calmodulin-dependent enzymes display similar cooperative activation kinetics.

We first compared the activation by Ca^{2+} of the calmodulin-dependent ATPase activities of human erythrocyte ghosts that had been prepared in various media, but which were all assayed in conditions of ionic strength and of ATP and Mg^{2+} concentration relatively similar to those in the cell. We observed that membranes isolated in 2.5 mM HEPES/1 mM EGTA at pH 7.0 exhibited the behaviour shown in Fig. 1a: activation occurred at lower Ca^{2+} concentrations than with other ghost preparations, and full activation occurred over a narrow range of Ca^{2+} concentrations. A rise from 10 to 90% of full activity required only a four- to fivefold increase in Ca^{2+} concentration, indicating that Ca^{2+} activated the ATPase in a highly cooperative manner (Hill coefficient 3.0 ± 0.4 , $n = 5$). In the presence of a high concentration of calmodulin, substantial activity was expressed at least down to 0.1 μM (Fig. 1a). As the concentration of calmodulin was reduced, however, the proportion of the total activity expressed at micromolar concentrations of Ca^{2+} fell (Fig. 1a, b). These experimental results are very similar to those reported by Klinger *et al.*⁷ but these authors interpreted their data in terms of classical Michaelis-Menten kinetics and thus, we feel, overlooked the potential importance of their observations.

If these results reflect the pattern of activation of the Ca^{2+} pump in the intact erythrocyte, into which Ca^{2+} leaks only very slowly, the ionized Ca^{2+} concentration within the cell would automatically be maintained below 0.1 μM , that is, at the expected level for a healthy erythrocyte. That this pattern may indeed be the correct one, rather than an experimental artefact, was indicated by comparative experiments (not shown) in which the Ca^{2+} dependence of the Ca^{2+} -ATPase was assessed in several membrane preparations, using both unmodified samples and samples treated with 0.1 mg per ml saponin^{7,8}; the preparations tested were the new HEPES/EGTA ghosts, whole

erythrocytes, ghosts prepared in 20 mM Tris/1 mM EGTA and ghosts prepared in imidazole/histidine/EGTA mixtures⁷. In the absence of saponin, the three latter preparations showed either no ATPase (cells) or activation that was much less sensitive to Ca^{2+} and calmodulin than in the HEPES/EGTA preparation, but when saponin was present (to ensure membrane permeability) they all behaved like the saponin-free HEPES/EGTA ghosts illustrated in Fig. 1. We therefore believe that much of the reported variability in Ca^{2+} and calmodulin sensitivity of previously described ghost preparations arose simply because they were partially (and variably) resealed, so that access of added Ca^{2+} , ATP or calmodulin to the ATPase was restricted.

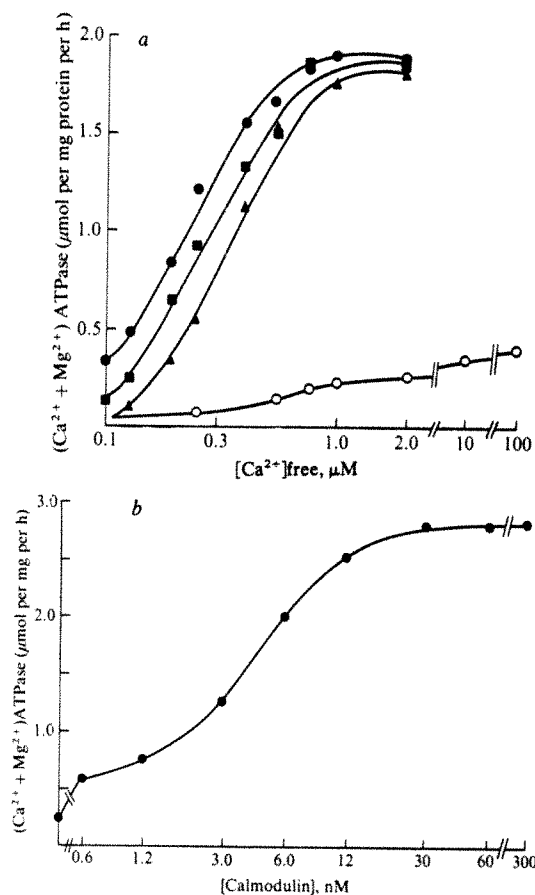


Fig. 1 Activation of $(\text{Ca}^{2+} + \text{Mg}^{2+})$ ATPase by *a*, Ca^{2+} in the absence (○) or presence of bovine brain calmodulin (▲, 56 nM; ■, 280 nM; ●, 5,600 nM), and *b*, calmodulin in the presence of a maximally stimulating concentration of Ca^{2+} (2 μM). Human blood was obtained from the Regional Blood Transfusion Laboratories within 7 days of collection. Erythrocytes were sedimented and washed three times in 0.15 M NaCl/1.5 mM HEPES, pH 7.2, at 4 °C to remove plasma and leukocytes. Cells were lysed in 10 vol of 2.5 mM HEPES/1 mM EGTA, pH 7.0. The membranes were sedimented at 15,000g for 15 min and then washed five times in the same medium. The isolated membranes were stored in this medium at 4 °C for up to 24 h before use. $(\text{Ca}^{2+} + \text{Mg}^{2+})$ ATPase activity was assayed for 30 min at 30 °C in 1 ml of 30 mM HEPES, 30 mM KCl, 80 mM NaCl, 4 mM MgCl_2 , 3 mM ATP, 0.1 mM ouabain, 1 mM EGTA, to which various concentrations of Ca^{2+} were added before adjustment of the pH to 7.0. Calmodulin was isolated from bovine brain¹⁶. Assays contained 0.15–0.2 mg of membrane protein. The ionized Ca^{2+} concentrations were calculated using $k_{\text{Ca}^{2+}\text{-EGTA}}$ of 11.0, as recommended by Owen¹⁷. Reactions were stopped by addition of cold trichloroacetic acid (final concentration 10% w/v) and P_i in the acid-soluble fraction measured¹⁸. Ca^{2+} -ATPase activity was calculated as total ATPase activity minus the activity expressed in the presence of 1 mM EGTA without added Ca^{2+} . The results shown in *a* and *b* are each taken from duplicate incubations in a single experiment, and are typical of results obtained in several experiments of fundamentally similar design. Results of the type shown were consistently obtained in many preparations of ghosts isolated in 2.5 mM HEPES, 1 mM EGTA, pH 7.0, but not in ghosts isolated in various other conditions.

Both we (Fig. 1) and others (for example, ref. 7) have noted that erythrocyte Ca^{2+} -ATPase can express its full activity at concentrations of calmodulin far lower than those present in the cell ($\sim 5 \mu\text{M}$)^{2,7}. Our results (closed symbols in Fig. 1a) and those of Klinger *et al.*⁷ also demonstrate that the calmodulin-dependent activation of the ATPase by Ca^{2+} retains its cooperativity at reduced calmodulin concentrations. However, as the calmodulin concentration is reduced, the apparent affinity for Ca^{2+} declines, suggesting that a function of the high, apparently 'excessive', calmodulin concentration that prevails within cells may be to maximize the sensitivity to Ca^{2+} of an already sensitive Ca^{2+} pump. Further studies using concentrations of Ca^{2+} and/or calmodulin that were inadequate for full activation gave complex results (not shown) that seem to rule out any explanation of activation involving simply the generation of a certain concentration of a particular Ca^{2+} -calmodulin complex as suggested by Cheung⁹; the relationship of these results to the mechanism of cooperative activation will be considered elsewhere.

At $\sim 1.0 \mu\text{M}$ Ca^{2+} , at which full activation of the ATPase was seen in the presence of calmodulin, the 'basal' (that is, calmodulin-independent) Ca^{2+} -ATPase activity of these ghosts was no more than 5% of their maximum activity, but this could be raised, in a relatively non-cooperative manner (open symbols in Fig. 1a), to $\sim 20\%$ of maximum activity at a Ca^{2+} concentration of $\sim 100 \mu\text{M}$. This activity persisted through extensive and varied washing procedures, as did a residual complement of $62 \pm 15 \text{ ng}$ ($n = 3$) of membrane-associated calmodulin per mg of membrane protein (as assayed by the activation of skeletal muscle myosin light chain kinase by heat-denatured erythrocyte ghosts; H. Dolbear and P.D., unpublished results). It therefore seems that soluble calmodulin must be available if the activation of the ATPase by Ca^{2+} is to show its normal cooperativity, but it remains possible that control by Ca^{2+} of the lower-affinity, 'calmodulin-independent' basal activity may still be mediated, in a different manner, by residual calmodulin that is tightly bound in the membrane preparations, as has been suggested previously^{7,10}. If this is correct, this residual calmodulin might be as tightly bound to the membrane ATPase as the calmodulin that is bound into the subunit structure of phosphorylase kinase.

Recent work has suggested that some Ca^{2+} -mobilizing hormones and neurotransmitters evoke intracellular responses by causing relatively small (for example, fivefold) changes in the cytosolic ionized Ca^{2+} concentration (for example, ref. 11). Such small changes in Ca^{2+} concentration should therefore be able to control the key intracellular enzymes responsible for these responses. At least two calmodulin-dependent Ca^{2+} -activated enzymes other than the Ca^{2+} -ATPase (myosin light chain kinase^{12,13} and cyclic nucleotide phosphodiesterase¹⁴) show highly cooperative activation by low concentrations of Ca^{2+} . Recent data suggest that the same may also apply to glycogen phosphorylase kinase when activated through its endogenous calmodulin subunit but not when activated by exogenous troponin I; this cooperativity was not, however, noted in the original paper¹⁵. It seems possible that such cooperativity may be an essential and general characteristic for the functioning of many enzymes that must respond to small changes in cytosol Ca^{2+} concentration. It will be interesting to determine whether the same type of mechanism underlies the cooperative activation by Ca^{2+} of all these calmodulin-dependent enzymes.

We thank the MRC for the award of a research studentship to P.D.

Received 4 November 1980; accepted 5 February 1981.

- Schatzmann, H. J. *Curr. Topics Membranes Trans.* **6**, 125-168 (1975).
- Jarrett, H. W. & Penniston, J. T. *Biochem. biophys. Res. Commun.* **77**, 1210-1216 (1977).
- Gopinath, R. M. & Vincenzi, F. F. *Biochem. biophys. Res. Commun.* **77**, 1203-1209 (1977).
- Scharff, O. & Foder, B. *Biochim. biophys. Acta* **509**, 67-77 (1978).
- Hanahan, D. J., Taverna, R. D., Flynn, D. D. & Ekholm, J. E. *Biochem. biophys. Res. Commun.* **84**, 1009-1015 (1978).
- Muallem, S. & Karlsh, S. J. D. *Biochim. biophys. Acta* **597**, 631-636 (1980).
- Klinger, R., Wetzker, R., Fleischer, I. & Frunder, H. *Cell Calcium* **1**, 229-240 (1980).
- Hanahan, D. J. & Ekholm, J. E. *Archs Biochem. Biophys.* **187**, 170-179 (1978).
- Cheung, W. Y. *Science* **207**, 19-27 (1980).
- Lynch, T. J. & Cheung, W. Y. *Archs Biochem. Biophys.* **194**, 165-170 (1979).

- Murphy, E., Coll, K., Rich, T. L. & Williamson, J. R. *J. biol. Chem.* **255**, 6600-6608 (1980).
- Nairn, A. C. & Perry, S. V. *Biochem. J.* **179**, 89-97 (1979).
- Stull, J. T. *Adv. Cyclic Nucleotide Res.* **13**, 39-93 (1980).
- Dedman, J. R., Potter, J., Jackson, R. L., Johnson, J. D. & Means, A. R. *J. biol. Chem.* **252**, 8415-8422 (1977).
- Cohen, P. *Eur. J. Biochem.* **111**, 563-574 (1980).
- Grand, R. J. A., Perry, S. V. & Weeks, R. A. *Biochem. J.* **177**, 521-529 (1979).
- Owen, J. D. *Biochim. biophys. Acta* **451**, 321-325 (1976).
- Baginski, E. S., Zak, B. & Foa, P. P. *Clin. Chem.* **13**, 326-332 (1967).

Occluded bound calcium on the phosphorylated sarcoplasmic transport ATPase

H. Takisawa & M. Makinose

Max-Planck-Institut für medizinische Forschung, Heidelberg, Abteilung Physiologie, 6900 Heidelberg, Jahnstrasse 29, FRG

The Ca^{2+} + Mg^{2+} -activated ATPase of the sarcoplasmic reticulum is responsible for the active Ca^{2+} transport of this membrane system¹, the key feature of which is the formation of an energy-rich phosphorylated transport enzyme (EP) and its conversion². To understand the Ca^{2+} -transport mechanism, it is essential to clarify the behaviour of this intermediate in relation to such ligands as ATP, ADP, Mg^{2+} and, particularly, Ca^{2+} . Recent kinetic studies on the phosphate turnover of this system suggested a relatively slow rate of Ca^{2+} dissociation from the phosphorylated enzyme^{3,4}, which possibly indicated Ca^{2+} binding in some occluded form with the intermediate. Here we report direct measurements of the binding and release of Ca^{2+} during phosphorylation of the sarcoplasmic transport enzyme. The results indicate an occlusion of the Ca^{2+} binding, accompanying an initial configurational change of the enzyme induced by the energy-rich phosphoryl transfer.

Figure 1 shows Scatchard plots of Ca^{2+} binding by the vesicular and solubilized sarcoplasmic transport ATPase in the absence of ATP. No essential difference could be observed between assays using 1 and 5 mM MgCl_2 . The curves for both preparations show a break. In the case of vesicular ATPase, the calcium affinities of the preparation, calculated from the slopes of the curves, were $k_{\text{dis}} \geq 1 \mu\text{M}$ and $\approx 10 \mu\text{M}$, which suggests negatively cooperating sites for Ca^{2+} binding^{5,6}. The maximal capacity for Ca^{2+} binding, obtained by extrapolation of the curves in Fig. 1 to the ordinate intercept, corresponds to 2.2-2.3 mol Ca^{2+} per mol of the ATPase phosphorylatable site in the vesicular preparation. Numerous kinetic analyses of the transient and/or steady state activity of the native or the partially purified vesicular sarcoplasmic ATPase have shown that one molecule of the transport ATPase is activated by the binding of two Ca^{2+} (ref. 7). The maximal Ca^{2+} binding capacity obtained from Fig. 1 includes, therefore, 0.2-0.3 mol of Ca^{2+} binding capacity per mol phosphorylatable site in excess. This could be due to the nonspecific Ca^{2+} binding capability of the preparation, for example, due to contamination of the Ca^{2+} -binding protein or of denatured ATPase protein.

All the ATPase preparations formed, with ATP, a maximum of 4-5 nmol phosphoenzyme per mg protein. Based on a molecular weight of 115,000 for the sarcoplasmic ATPase⁸, this indicates that about half of the vesicular or solubilized ATPase preparations consist of enzymatically inert proteins—that is, mainly denatured ATPase protein.

However, the solubilized ATPase protein shows a maximal Ca^{2+} binding capacity of ~ 1.4 mol per mol of phosphorylatable site. As both the ATPase preparations show the same specific activity, it is reasonable to assume that they both contain the same amount of inert protein. Thus, the subtraction of nonspecific Ca^{2+} binding capacity found in the vesicular preparation gave a maximal Ca^{2+} specific binding capacity of the solubilized ATPase which was almost half that of the former preparation. A break in the Scatchard plots of solubilized

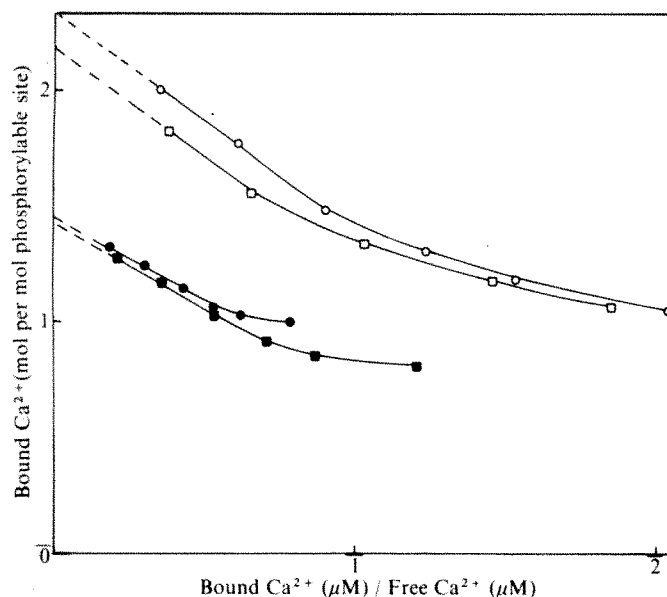


Fig. 1 Scatchard plots of Ca^{2+} binding by the vesicular and solubilized Ca^{2+} -transport ATPase. Binding of Ca^{2+} to the ATPase was determined by the flow-dialysis method of Colowick and Womack¹⁰ in solutions containing 0.4 M KCl, 0.3 M sucrose, 1 (○, ●) or 5 mM (□, ■) MgCl_2 and 50 mM Tris-TES (tri-methylamino ethanesulphonic acid), pH 8, at 5°C. The concentration of the enzymes was adjusted to 3 mg ml⁻¹ and, in the case of solubilized ATPase (●, ■), 3.5 mg ml⁻¹ deoxycholate (DOC) was included in the assay. Vesicular ATPase (partially purified sarcoplasmic reticulum vesicles: ○, □) was prepared by the method of Meissner *et al.*¹¹. To obtain the solubilized ATPase, 3.5 mg ml⁻¹ of the sarcoplasmic protein was solubilized with DOC (3.5 mg ml⁻¹) in the presence of 0.4 M KCl, 0.3 M sucrose and 50 mM Tris-TES, pH 8, at 5°C¹². After pelleting of the non-solubilized proteins (105,000g for 50 min), the supernatant contained ~3 mg of protein and 3.5 mg of DOC per ml. About 6 nmol Ca^{2+} per mg protein (obtained by atomic absorption spectrometry, Perkin-Elmer Model 4000) constituted Ca^{2+} contamination in both enzyme preparations. The measurement of Ca^{2+} binding was started by the addition of 4 nmol (3.2 μCi) ⁴⁵Ca to 1 ml of the enzyme solution. The Ca^{2+} concentration was increased stepwise up to 2 mM by the addition of 'cold' CaCl_2 . The ordinate is the amount of Ca^{2+} bound per mol of phosphorylatable site of the preparation used. The maximum amount of phosphoprotein was estimated in an assay containing 3 mM CaCl_2 , 1 mM MgCl_2 and 100 μM ATP, as 4–5 nmol EP per mg protein for both the vesicular and solubilized ATPase.

ATPase (Fig. 1) is presumably due to the nonspecific bound Ca^{2+} capacity. Apparently, the solubilization reduces the affinity of one of the two sites of the enzyme for Ca^{2+} .

This raises the question of whether the solubilized sarcoplasmic ATPase requires one or two Ca^{2+} for its activation. Because the rate of the EP formation of the solubilized ATPase is much slower than that of the vesicular ATPase (data not shown), the time course of EP formation can be easily followed. Figure 2 shows the change of EP level, extra Ca^{2+} binding and P_i liberation during transient EP formation. It is evident that the extra binding of Ca^{2+} was accompanied by EP formation and the amount of additional Ca^{2+} bound to the ATPase was almost equal to the amount of EP formed. No correlation was found between the time course of additional Ca^{2+} binding and P_i liberation. Apparently, two Ca^{2+} are also required for the activation of the solubilized sarcoplasmic ATPase. The reaction steps for the activation of the solubilized ATPase can thus be formulated:

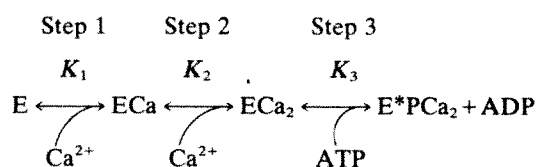


Table 1 Retention of bound calcium on the phosphorylated sarcoplasmic transport enzyme

ATPase preparation	Addition	Effluent contents		
		Calcium* (nmol)	Phosphorylatable site (nmol)	Mol bound Ca^{2+} per mol phosphorylatable site
Vesicular	None	0.180 ± 0.005	0.244 ± 0.004	0.739 ± 0.010
	None	0.099 ± 0.005	0.306 ± 0.004	0.324 ± 0.015
Solubilized	ATP (100 μM)	0.371 ± 0.011	0.259 ± 0.019	1.437 ± 0.076
	AMPPCP (200 μM)	0.091 ± 0.007	0.281 ± 0.007	0.322 ± 0.024

A small Sephadex G-50 column (4.7 × 57 mm) prepared according to Penefsky⁹ was equilibrated with a solution containing 0.4 M KCl, 0.3 M sucrose and 50 mM Tris-TES, pH 8. The enzymes (3 mg protein ml⁻¹), were incubated with ⁴⁵CaCl₂ (1 mM) in the presence or absence of ATP (100 μM) or AMPPCP (200 μM) in a solution containing 0.4 M KCl, 0.3 M sucrose, 80 μM MgCl_2 and 50 mM Tris-TES, pH 8 at 0°C for 2 min. An aliquot of the assay (50 μl) was transferred to the top of the column and centrifuged immediately (150g for 2 min at 5°C) and the effluent calcium and protein content were measured. The enzyme preparations showed the maximal phosphorylatable site number of 5.1 nmol per mg protein using the method described in Fig. 1 legend, with no significant difference between the vesicular and the solubilized enzyme. Values are the means of four determinations ± s.e.m.

* The calcium content of the effluent in the control experiment containing no protein was <0.002 nmol.

The solubilizing procedure affects the ATPase by reducing its affinity for the second Ca^{2+} (K_2), so that, at the concentration of Ca^{2+} which is usually sufficient to activate the transport enzyme (for example, 90 μM as in Fig. 2), one solubilized ATPase molecule binds about one Ca^{2+} . If ATP is added, step 3 of the reaction is activated and the enzyme is accumulated in the form of E*PCa_2 , which retains two Ca^{2+} per molecule. The binding state of the Ca^{2+} with this enzyme species should be very different from that of the unphosphorylated enzyme species.

In the experiments shown in Table 1, the ATPase preparations, after incubation in appropriate conditions, were centrifuged through a small Sephadex column to rapidly remove the incubation medium (centrifuge column method⁹). As the enzyme is exposed to the ligand free-solution in the column

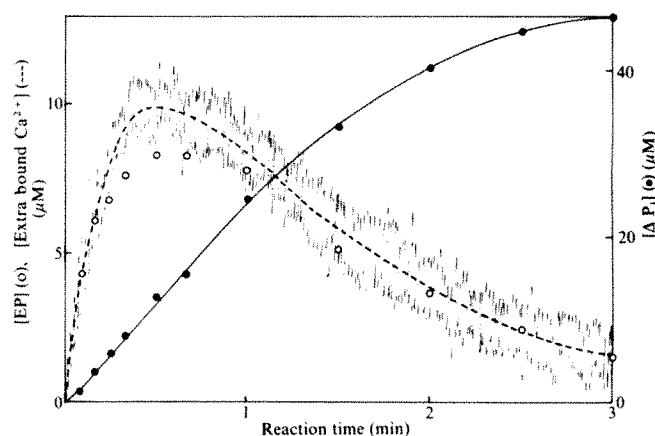


Fig. 2 Extra Ca^{2+} binding to the solubilized sarcoplasmic ATPase during EP formation. The reaction was started by adding 50 μM [γ -³²P]ATP to the assay containing 3 mg ml⁻¹ solubilized ATPase protein, 3.5 mg ml⁻¹ DOC, 0.4 M KCl, 0.3 M sucrose, 90 μM CaCl_2 , 80 μM MgCl_2 and 50 mM Tris-TES pH 8.0, at 5°C and terminated after the appropriate time by addition of 4% trichloroacetic acid containing 0.1 mM ATP and 0.2 mM P_i . EP formation and P_i liberation were measured as described elsewhere^{13,14}. Change in the concentration of Ca^{2+} was measured in the same assay containing 100 μM Murexide¹⁵. The reaction was started with quick addition (within 1 s) of 50 μM cold ATP and the absorbance change of the dye at 507/540 nm was registered on a Aminco DW-2 spectrophotometer. The addition of ATP caused an immediate change in absorbance corresponding to the decrease in Ca^{2+} concentration of ~2 μM due to formation of Ca -ATP complex. This deviation was subtracted from the original trace.

during the separation time, the bound ligands, particularly the rapidly exchangeable ligands, are partially removed from the enzyme protein. When the vesicular ATPase was preincubated in solution containing 1 mM Ca^{2+} , 0.74 mol of Ca^{2+} was bound per mol active enzyme after centrifugation instead of the expected 2 mol. In the case of solubilized ATPase, 0.32 mol of Ca^{2+} was bound per mol of ATPase instead of the expected 1 mol. Apparently, about two thirds of the rapidly exchangeable bound Ca^{2+} is removed from the ATPase protein in this procedure. When the solubilized ATPase was incubated with 100 μM ATP and 1 mM Ca^{2+} , 1.44 mol of Ca^{2+} were bound per mol active enzyme in the effluent. In a parallel assay with 'hot' ATP and 'cold' Ca^{2+} , in the effluent, 32% of the phosphorylatable sites of the enzyme were unphosphorylated and 68% were phosphorylated. Therefore, 1 mol of the active ATPase in these effluents retains 0.1 mol (32% of 0.32 mol) Ca^{2+} , which represents the unphosphorylated state and the remaining 1.34

mol Ca^{2+} represents the phosphorylated state (68% of 1 mol). Thus, even after column centrifugation, 2 mol of Ca^{2+} are retained by 1 mol phosphorylated enzyme. Addition of 200 μM AMPPCP (non-hydrolysable ATP analogue) had no effect on the Ca^{2+} -binding of the solubilized ATPase.

The condition for EP formation in which bound Ca^{2+} cannot be removed is low Mg^{2+} concentration (80 μM). When the experiment in Fig. 2 was done using 5 mM Mg^{2+} , addition of ATP caused a transient increase in Ca^{2+} concentration in the assay, as observed previously by Ikemoto⁵. A high Mg^{2+} concentration presumably effects the further conversion of the phosphoprotein, resulting in Ca^{2+} liberation and dephosphorylation. The occlusion of the bound calcium on the transport ATPase molecule is coupled with an initial configurational change of the enzyme molecule caused by phosphorylation of the enzyme protein, but not by formation of the substrate-enzyme complex.

Received 28 July 1980; accepted 2 January 1981.

1. Hasselbach, W. & Makinose, M. *Biochem. Z.* **339**, 99–111 (1963).
2. Tonomura, Y. *Muscle Proteins, Muscle Contraction and Cation Transport* (University of Tokyo and University Park Press, Tokyo and Baltimore, 1972).
3. Takakuwa, Y. & Kanazawa, T. *Biochem. biophys. Res. Commun.* **88**, 1209–1216 (1979).
4. Takisawa, H. & Tonomura, Y. *J. Biochem., Tokyo* **86**, 425–441 (1979).
5. Ikemoto, N. *J. Biol. Chem.* **250**, 7219–7224 (1975).
6. Yates, D. W. & Duance, V. C. *Biochem. J.* **159**, 719–728 (1976).
7. Yamamoto, T., Takisawa, H. & Tonomura, Y. in *Current Topics in Bioenergetics* Vol. 9 (ed. Sanadi, D. R.) (Academic, New York, 1979).

8. Le Maire, M., Jørgensen, K. E., Røigaard-Petersen, H. & Møller, J. V. *Biochemistry* **15**, 5805–5812 (1976).
9. Penefsky, H. S. *J. Biol. Chem.* **252**, 2891–2899 (1977).
10. Colowick, S. P. & Womack, F. C. *J. Biol. Chem.* **244**, 774–777 (1969).
11. Meissner, G., Conner, G. E. & Fleischer, S. *Biochim. biophys. Acta* **298**, 246–269 (1973).
12. Jørgensen, K. E., Lind, K. E., Røigaard-Petersen, H. & Møller, J. V. *Biochem. J.* **169**, 489–498 (1978).
13. Makinose, M. *Eur. J. Biochem.* **10**, 74–82 (1969).
14. Takisawa, H. & Tonomura, Y. *J. Biochem., Tokyo* **83**, 1275–1284 (1978).
15. Scarpa, A. *Meth. Enzym.* **24**, 343–351 (1972).

Direct observation by NMR of two coexisting conformations of an enzyme–ligand complex in solution

A. Gronenborn, B. Birdsall, E. I. Hyde, G. C. K. Roberts*, J. Feeney & A. S. V. Burgen

Division of Molecular Pharmacology, National Institute for Medical Research, London NW7 1AA, UK

Dihydrofolate reductase is the site of action of the 'anti-folate' drugs, such as methotrexate, pyrimethamine and trimethoprim, which act as competitive inhibitors of the enzyme^{1–3}. These inhibitors bind to the *Lactobacillus casei* enzyme cooperatively with the coenzyme NADP(H), which can increase inhibitor binding by as much as 700-fold⁴. The existence of conformational changes accompanying ligand binding and involved in inhibitor–coenzyme cooperativity has been inferred from NMR experiments^{4–9}. It seems likely that these conformational changes are important in determining the specificity of the enzyme for its ligands^{4,10,11}. We now report the direct observation by NMR of two distinct conformations of the enzyme–trimethoprim–NADP⁺ ternary complex in solution.

Figure 1 compares the histidine C¹–H resonances of this complex with those of the enzyme–methotrexate–NADP⁺ complex. The resonances of all seven histidine residues¹², labelled A–G, are clearly resolved in these conditions. Their positions are closely similar in the two complexes (for a detailed comparison, see ref. 13), with the exception of resonance F. This has been assigned to His 28 (ref. 14), which is seen in the crystal structure¹⁵ of the enzyme–methotrexate–NADPH complex to interact with the γ -carboxyl of the glutamate moiety of methotrexate. As there is no corresponding group in trimethoprim, the pK of His 28 is ~ 1.3 units lower in complexes with trimethoprim than in those with methotrexate^{13,16}.

The most striking difference between the two spectra in Fig. 1, however, is that in the spectrum of the trimethoprim complex resonances E (tentatively assigned to His 18 (refs 13, 14)) and F are split into two peaks, each of intensity corresponding to approximately half a proton (compare resonances A, B and D).

The separation between these two components of resonances E and F, measured in hertz, is directly proportional to magnetic field strength and therefore represents a chemical shift difference. The spectra thus show that the enzyme–trimethoprim–NADP⁺ complex exists in two forms which differ in the environments of two histidine residues, His 28 and, probably, His 18. The approximately equal intensity of the two components of the histidine resonances indicates that these two forms are present in approximately equal amounts.

If these two forms interconvert with one another, they must represent different conformational states of the complex. The observations of two distinct signals separated by only ~ 15 Hz (at 270 MHz and 5 °C) implies that such an interconversion must be slow. We have demonstrated that interconversion occurs and measured its rate by determining the temperature dependence of the line shape of histidine resonances E and F over the range 0–50 °C at both 270 and 470 MHz. The results at 470 MHz are shown in Fig. 2.

As the temperature is increased, the two components of each histidine resonance broaden, and then coalesce into a single line which sharpens as the temperature is increased further. This behaviour is that expected for a progressively increasing rate of interconversion between the two states represented by the two components of the absorption band. (The progressive downfield shift of the resonances with increasing temperature is due to the temperature dependence of the pK values of the histidine residues.) These changes in line shape, and particularly the broadening of the resonances with increasing temperature at low temperature, cannot be accounted for by any model in which there is no interconversion between the two states, thus eliminating the possibility that they correspond to two hitherto undetected isoenzymes of the kind observed in *Escherichia coli* RT500 (ref. 17).

The observed line shapes can be satisfactorily simulated, as indicated in Fig. 2, on the basis of a model involving interconversion between just two states of the protein. These simulations allow us to estimate the rate of interconversion at each temperature, taking the chemical shift difference between the two states as that measured at 5 °C (0.044 p.p.m. for resonance F, 0.063 p.p.m. for resonance E). The estimated rates range from 2 s^{–1} at 26 °C to 80 s^{–1} at 51 °C. Four independent estimates of the rate can be made at each temperature from simulation of resonance E and resonance F each at 270 and 470 MHz. These four estimates always agreed within $\pm 15\%$,

* To whom correspondence should be addressed.

indicating that the two-state model of the interconversion process is adequate. The fact that the two components of resonance F coalesce at a lower temperature than those of resonance E can be simply explained by the smaller chemical shift difference between the two states in the former case. The activation parameters for the interconversion, calculated from the temperature dependence of the estimated rate constants, are: $\Delta H^\ddagger = 25(\pm 3)$ kcal mol⁻¹, $\Delta S^\ddagger = 29(\pm 8)$ cal K⁻¹ mol⁻¹. There is no detectable change in the relative proportions of the two states over the temperature range studied.

Note that the splitting of the two histidine resonances is not observed in either the enzyme-trimethoprim or the enzyme-NADP⁺ binary complexes. The observation of two coexisting conformations of the ternary complex but not of either binary complex is clear evidence for the existence of conformational changes specific to the ternary complex. These may be involved in the observed cooperativity in binding between NADP⁺ and trimethoprim.

Examination of the three-dimensional structure of the enzyme-methotrexate-NADPH complex¹⁵ shows that the backbone carbonyl group of His 18 forms a hydrogen bond to the 3'-hydroxyl of the nicotinamide ribose of the bound coenzyme. His 28, on the other hand, is farther both from the bound coenzyme and from trimethoprim (bound in the conformation deduced from ¹H NMR data¹⁸). We have recently shown¹⁹ that the interaction of the nicotinamide ring of the coenzyme with the enzyme and the conformation of the bound coenzyme both differ between the two conformations of the ternary complex. The two environments of His 18 may thus reflect the two modes of interaction of the coenzyme. The effect on His 28, however, indicates that the conformational differences between the two states of the complex extend beyond the residues in direct contact with the bound ligands.

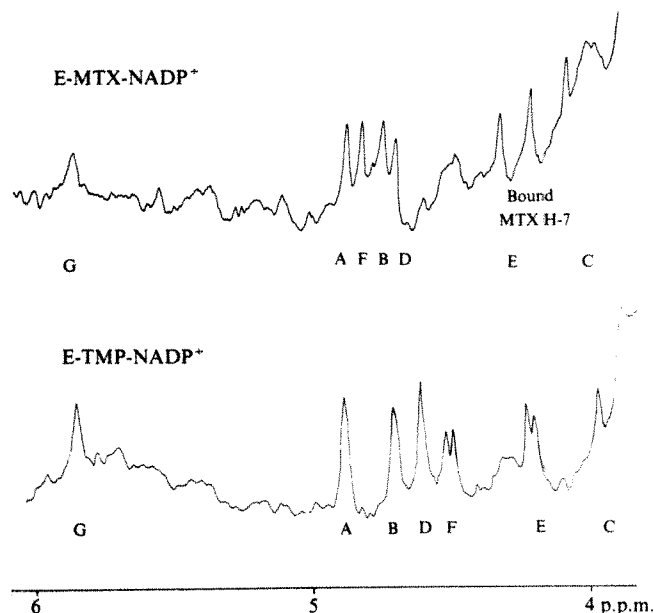


Fig. 1 Histidine C^{ε1}-proton resonance region of the ¹H NMR spectrum of *L. casei* dihydrofolate reductase in its complexes with methotrexate (MTX) and NADP⁺ (top) or trimethoprim (TMP) and NADP⁺ (bottom). The resonances of the seven histidine residues are labelled A–G. The spectra were obtained at 270 MHz using a Bruker WH270 spectrometer in the Fourier transform mode. The samples consisted of ~1 mM enzyme in 0.35 ml² H₂O containing 50 mM potassium phosphate, pH (meter reading) 6.5, 500 mM KCl. Approximately 2,000 transients obtained using quadrature detection, with 8,192 data points for a spectral width of 4.2 kHz, were averaged for each spectrum. Before Fourier transformation, the free induction decay was multiplied by an exponential equivalent to a line broadening of 2 Hz. Chemical shifts are expressed relative to internal (1 mM) dioxan (3.71 p.p.m. downfield from dimethylsilapentane-5-sulphonate).

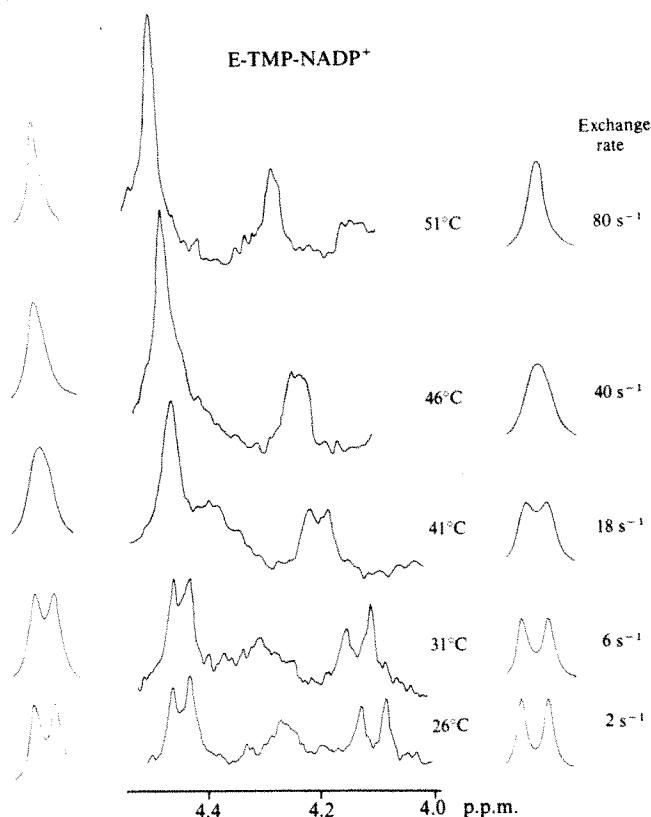


Fig. 2 The histidine C^{ε1}-proton resonances E (on the right) and F of the dihydrofolate reductase-trimethoprim-NADP⁺ complex as a function of temperature. The line shapes simulated using the equation for a two-site exchange process, with the chemical shift difference measured at 5 °C and the rate constants shown (chosen for the best visual fit to the experimental spectra), are shown at either side. The experimental spectra were obtained as described in Fig. 1 legend, but using the 470-MHz spectrometer of the Oxford Enzyme Group.

The coexistence of two or more conformational states of an enzyme complex has often been postulated, or inferred from indirect evidence, in this and other systems. In the present case, however, we are able to observe two coexisting conformational states directly in solution, measure their rate of interconversion and begin to characterize the structural differences between them.

We thank Gill Ostler and John McCormick for technical assistance and the Oxford Enzyme Group, particularly Iain Campbell and Peter Stiles, for the use of their 470-MHz NMR spectrometer. E.I.H. is supported by a MRC Research studentship.

Received 1 September 1980; accepted 5 February 1981.

- Blakley, R. R. L. *The Biochemistry of Folic Acid and Related Pteridines* (North-Holland, Amsterdam, 1969).
- Kisliuk, R. L. & Brown, G. M. (eds) *Chemistry and Biology of Pteridines* (Elsevier, Amsterdam, 1979).
- Burchall, J. J. & Hitchings, G. H. *Molec. Pharmac.* **1**, 126–136 (1965).
- Birdsall, B., Burgen, A. S. V. & Roberts, G. C. K. *Biochemistry* **19**, 3723–3731 (1980).
- Hyde, E. I., Birdsall, B., Roberts, G. C. K., Feeney, J. & Burgen, A. S. V. *Biochemistry* **19**, 3738–3746 (1980).
- Hyde, E. I., Birdsall, B., Roberts, G. C. K., Feeney, J. & Burgen, A. S. V. *Biochemistry* **19**, 3746–3754 (1980).
- Feeney, J. *et al. Proc. R. Soc. B* **196**, 267–290 (1977).
- Roberts, G. C. K. *et al. in NMR in Biology* (eds Dwek, R. A., Campbell, I. D., Richards, R. E. & Williams, R. J. P.) **95** (Academic, London, 1977).
- Feeney, J. *et al. Biochemistry* **19**, 2466–2472 (1980).
- Roberts, G. C. K. *Ciba Fdn Symp.* **60**, 89–103 (1978).
- Birdsall, B., Burgen, A. S. V., de Miranda, J. P. R. & Roberts, G. C. K. *Biochemistry* **17**, 2102–2110 (1978).
- Bitar, K. G. *et al. FEBS Lett.* **80**, 119–122 (1977).
- Gronenborn, A. *et al. Biochemistry* (in the press).
- Wyeth, P. *et al. Biochemistry* **19**, 2608–2615 (1980).
- Matthews, D. A. *et al. J. Biol. Chem.* **253**, 6946–6954 (1978).
- Birdsall, B., Griffiths, D. V., Roberts, G. C. K., Feeney, J. & Burgen, A. S. V. *Proc. R. Soc. B* **196**, 251–265 (1977).
- Baccanari, D. P., Averett, D., Briggs, C. & Burchall, J. *Biochemistry* **16**, 3566–3572 (1977).
- Cayley, P. J. *et al. Biochemistry* **18**, 3886–3895 (1979).
- Gronenborn, A. *et al. Molec. Pharmac.* (in the press).

BOOK REVIEWS

Computers born of crisis

S.J. Goldsack

A History of Computing in the Twentieth Century. Edited by N. Metropolis *et al.* Pp.688. (Academic: 1980.) \$29.50, £16.60.
Early British Computers. By S. Lavington. Pp.139. (Manchester University Press: 1980.) Hbk £7.95; pbk £3.95.

SEPARATELY, these books would both be well worth reading by anyone with an interest in the origins of computing as we know it today. Together I found them fascinating. Both volumes claim to be of interest to the non-specialist reader, as well as to the technically aware. I would judge that this claim is reasonably justified; however, I think that an understanding of the technical background would certainly help.

The book by Metropolis *et al.* is the more important. In June 1976, a conference on the history of computing was held in Los Alamos. Its purpose was to bring together as many as possible of the pioneers who created the modern computer, to tell the story of their work. The timing was none too early; already important figures such as J. von Neumann, A. Turing and F.C. Williams, whose presence would have added greatly to the proceedings, had died. Their places were taken by colleagues who had shared in their work. Speakers described the early work on the design and use of computers in America, England, Germany, Russia and Japan. The book is essentially the written record of that conference.

Many of the basic concepts involved in computing were understood before the outbreak of war in 1939. However, the urgent need for tools of computation in connection with war work, and the great technical advances in electronics which were brought about by the efforts in radar and other fields, produced circumstances in which the construction of electronic computers could prosper.

It is commonplace to recognize how the threat of war in the 1930s, and the war effort that followed, fostered technical advances in many fields. The participants in these efforts experienced a uniquely exciting period in their lives. The intrinsic scientific interest of the work in which they were involved, the urgency with which the research was pursued, and, perhaps for the first time, a political commitment by the governments involved to ensure that the resources required were as far as possible made available, combined to create an atmosphere which had not before been



M.J. Lanigan and Tom Kilburn with a section of the transistorized MUSE computer in 1959. This prototype was a forerunner of the Ferranti ATLAS, inaugurated by Sir John Cockcroft in 1962, which incorporated several new concepts in computer design.

encountered. After 1945, those involved returned to the universities, or to newly established research centres, fired with enthusiasm to continue the work which they had started in the urgent circumstances of international crisis. As a student in the immediate post-war period, it was impossible not to sense the enthusiasm which our teachers brought to their work.

The authors of this book were not primarily writing memoirs of their experiences; yet the effect of the collection of accounts by all the parties participating in work relating to electronic computers about how progress was achieved in the different centres, and the interrelation between them, creates a strong impression of the vitality of the period.

It would not be realistic to single out particular papers as worthy of special notice, as the interest of the book lies in the way in which the contributions of the various authors combine to create a complete picture. Inevitably, however, there is special interest in those papers describing work done in conditions of secrecy during the war, and of which these accounts are among the first publicly available. They include, for example, the description of the Colossus machines, used to decipher German codes by the British Secret Service — oddly, this work was sponsored by the Foreign Office! I personally learned much of which I was

ignorant, or only dimly aware. I had failed to appreciate, for example, how widely the Williams storage tube was used on both sides of the Atlantic for a period before the invention of magnetic core stores.

The volume addresses the history of computing. Though hardware has pride of place, the developments in programming and programming languages are included. It would be impossible to appreciate the development of computer architecture, without understanding the need for better facilities for the programmer.

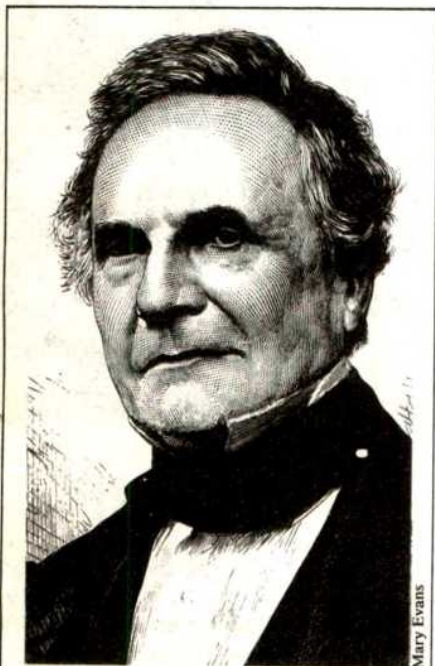
The date chosen, tacitly, to determine the end of history and the start of the modern era was rather arbitrary. Historical computers, it seems, used valves: transistors belong to the modern era. The computers of the second generation, which saw the development of operating systems and complex software, were not covered. Perhaps a future conference should address the history of programming and software; it is doubtful, however, whether it would reach the same level of interest for the general reader as is achieved in this book.

It is not surprising, perhaps, that Lavington's volume should appear at about the same time as the other. Lavington presented the work of the Manchester group under F.C. Williams at Los Alamos, and, no doubt, his book was inspired in part by the research involved in

preparing his contribution to the larger work. Though it has less scope and significance, it does include an account of the ATLAS Computer which was largely developed at Manchester and which introduced the concept of virtual storage in the early 1960s.

Lavington's early chapters, which attempt to give in a few pages a technical account of the early developments, are rather shaky. However, the style quickly settles down to an interesting description which reinforces and extends the material of the larger book. □

S.J. Goldsack is Professor in the Department of Computing at Imperial College, University of London.



Charles Babbage, 1792–1871, who devised the basic principles of modern computers.



Alan Turing, a leading figure in the early days of computer design. He died in 1954.

Computing and optics in context

P.W. Hawkes

The Computer in Optical Research. Edited by B.R. Frieden. Pp.371. (Springer-Verlag: 1980.) DM 98, \$57.90.

SOME topics seem so naturally adapted to a multi-author series such as this that when the corresponding volume does appear it is hard to realize that it has no rivals. Frieden's collection is a good example: most of the material is available elsewhere, scattered throughout the literature though not necessarily that of optics, but how very much more useful and usable it becomes, gathered into a single volume, put in context and with a recognizable pattern imposed on it.

Five subject-areas are explored. First, R. Barakat describes the calculation of the various types of integrals that arise in diffraction theory; his aim is to draw attention to the methods that have been developed by numerical analysts, with which specialists in other fields are all too often unfamiliar. This is followed by a long and extremely important chapter by Frieden on probability and statistics. This contains a formal but highly readable account of those aspects of the subjects that Frieden believes to be valuable in optics. After arguing persuasively that such methods are likely to become all-pervasive in this field, he regrets that progress is slow since "most optical workers simply do not know the methods" even though "first applications in optics have been highly successful". He claims, quite justifiably, that his chapter is the "first exposition on probability and statistics written specifically from the

optical perspective and for the optical community".

The remainder of the book is less revolutionary though of considerable interest. A.K. Rigler and R.J. Pegis give a knowledgeable account of optimization methods, L. Mertz examines computers and optical astronomy and W.J. Dallas surveys the computer generation of holograms. Every chapter is filled with practical information, illustrated with numerous examples, and it is clear either that Frieden has submitted the chapters to thorough editing to ensure uniformity of tone or that he chose his contributors very skilfully. Furthermore, he has written an introductory chapter that adds substantially to the usefulness and readability of the remainder. In this, he writes what amounts to a long essay-review of the book, explaining informally how the individual chapters are organized and giving his own opinions of the place of particular approaches in optics. In addition he draws attention, sometimes anecdotally, to other topics that could have been included, or are at too preliminary a stage of development to appear in the main text; Fienup's modified version of the Gerchberg-Saxton algorithm is an example of this.

It is not often that the multi-author formula is entirely satisfactory but here it works to perfection. The best way of recommending this wholly excellent book is to say that I cannot understand how we have been managing without it. □

P.W. Hawkes is Maître de Recherches at the Laboratoire d'Optique Electronique du CNRS, Toulouse.

The double layer in electrochemistry

Andrew Hamnett

The Double Layer. Comprehensive Treatise on Electrochemistry, Vol.1. Edited by J. O'M. Bockris, B.E. Conway and E. Yeager. Pp.453. (Plenum: 1980.) \$49.50, £31.19.

THE Plenum Press have published a remarkable series of books on electrochemistry in recent years, and the *Comprehensive Treatise on Electrochemistry* is their latest and most ambitious venture. This first volume contains nine articles on aspects of the double layer written by such well-known electrochemists as Parsons, Bockris, Frumkin and Damaskin. Inevitably, these articles make rather varied intellectual demands on the reader, from the considerable in the case of Parsons' introductory article on thermodynamics, to the pedagogically straightforward

account by Reeves of the double layer without specific adsorption.

Parsons' account of the Gibbs adsorption isotherm is rightly given the task of introducing this volume. It is an essay of considerable erudition, though the development of the basic equations is very abstract. Some didactic edge is therefore lost in the early part of the essay and only recovered when the author considers some specific examples of electrode configuration.

The two articles by Reeves and by Habib and Bockris constitute the central part of the book. They are both clear and well written, though I would have preferred a more thorough and detailed account of the recent developments in double-layer theory described in Reeves' article. The essay by Bockris and Habib on specific ionic adsorption is a thorough account of a

highly controversial area; indeed, some of the flavour of this controversy is evident in the article, lending it at times an almost polemical tone. Equally clear is the discussion by Damaskin and Kazarimov on the adsorption of organic molecules onto electrodes; from the pedagogical point of view, this is the most successful article in the book.

Frumkin and co-authors present a detailed account of the concept of potential of zero charge. This topic has attracted several recent reviews and a book (by Frumkin); the present article is in fact a reprise of that book and will be welcomed therefore in the West. The importance of the concept of pzc is amply illustrated in the

present review and my only criticism is that there remain some problems in notation, both between this article and others in the volume, and indeed between text and figure captions.

The articles by Pleskov, Boguslavsky and Hunter are devoted to various aspects of the non-metal/electrolyte interface. Pleskov's essay reviews the double-layer on classical semiconductors and is an able summary of the famous book by Myamlin and Pleskov (*Electrochemistry of Semiconductors*; Plenum, 1967), but I was disappointed to find so few references to work of the past decade, especially given the explosion of interest that has occurred in this field. Boguslavsky has provided a

welcome if rather tentative account of the insulator/electrolyte interface and Hunter a brief but most useful description of colloids.

The present volume contains some essays of great value; in others the development is rather eclectic and the overall result is somewhat uneven. Nevertheless, the first volume of this series will undoubtedly become a standard work. Nothing quite so detailed has been attempted before in electrochemistry, and firmer editorial control should ensure that subsequent volumes will more than fulfil its promise. □

Andrew Hamnett is University Lecturer in Inorganic Chemistry and Fellow of St Catherine's College, Oxford.

Population genetics with age — not forgetting the genes

Michael Bulmer

Evolution in Age-Structured Populations. By Brian Charlesworth. Pp.300. (Cambridge University Press: 1980.) Hbk £18, \$44.50; pbk £5.95, \$13.50.

THIS is the first book in a new series called *Cambridge Studies in Mathematical Biology*, the aim of which is that "each book will be self-contained, covering necessary background theory, but without allowing the mathematical details to obscure biological relevance". The book under review admirably fulfils these aims.

The classical theory of population genetics assumes a life-history with discrete, non-overlapping generations in which all individuals in the population at a particular time are the same age. This model is appropriate for annual plants and insects, but not for a large number of species (most vertebrates, for example) with overlapping generations. In the latter situation, the population at a particular time consists of individuals of different ages with fecundities and survival probabilities dependent on their age, and a proper theory of its population genetics should take its age structure into account. In particular, it is clearly impossible to consider how natural selection moulds the way in which fecundity and survival depend on age without considering age structure, so that recent interest in the evolution of life-history strategies has stimulated theoreticians to extend classical population genetics theory to include age-structure as a variable. Brian Charlesworth reviews and summarizes this work, much of which is his own.

Five chapters develop the subject in a logical way. The first introduces the relevant demographic theory, and the next discusses population genetics in age-structured populations without selection (Hardy-Weinberg and linkage equilibrium, and genetic drift in finite populations); of particular value is the

description of the elegant work of Hajnal on the frequencies of consanguineous matings in an isolated population.

Chapters 3 and 4 form the core of the book and describe the effect of selection in an age-structured population. The author remarks that important work on this subject was published by Haldane and Norton in the late 1920s, but that further work was inhibited by the introduction by Fisher of his "Malthusian parameter" method in 1930, which apparently provided a simple and elegant answer to the problem. However, it has been recognized in the past ten years that this method is defective because a particular genotype cannot have a fixed Malthusian parameter when the genotypic composition of the population is changing under selection; there is no short-cut to the methodology introduced by Haldane and Norton.

Finally, Charlesworth discusses the evolution of life-history strategies, that is to say the evolution of senescence and of reproductive patterns. This is the most interesting area of application of the ideas developed in the book; the author's general conclusion is that

although qualitative agreement between theoretical expectation and observed differences in life-histories between and within species is often good, no studies of natural populations have really been capable of demonstrating close quantitative agreement between a real and an optimal life-history.

Few would disagree with this comment.

This is an important book which can be recommended to anyone with an interest in population genetics and evolutionary biology. The argument is presented clearly from first principles; the author does not belong to the hand-waving school of thought which believes that, when the going gets difficult, it is possible to do population genetics without genes. The mathematical level is reasonably straightforward, but it is at times difficult to see the

wood for the trees; a reader might do well to make a rapid first reading before studying the details.

Incidentally, it should be noted that Bernardelli is throughout mis-spelt "Bernadelli"; this is probably a copying error from a secondary source — a classical deletion mutation. □

Michael Bulmer is a Lecturer in Biomathematics at Oxford University.

SCIENTIFIC BOOKSHOP

H.K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.50. (Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

**H.K. LEWIS
& Co. Ltd.**
136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavit,
London, WC1E 6BS."

Circle No.26 on Reader Enquiry Card.

ISSN-0275-4207

THE BIOTECHNOLOGY NEWS SOURCE™



BioEngineering News -
the *only* biweekly
newsletter covering the
genetic-engineering
recombinant-DNA,
monoclonal-antibody,
and cell-culture
industries.

24 issues — \$236.00 (U.S.)
(\$195.00 in U.S. & Canada)

Subscriptions sent airmail.

Write: **BioEngineering News**, Suite
304, 109 Minna St., San Francisco,
CA 94105. Tel: (415) 332-0508.
Telex: 171463-HQLARK.

Special Reports Available: (1) Ge-
netic Engineering Industry, (2) Hy-
bridoma & Cell Fusion, (3) Auto-
mated DNA Synthesizers. Price:
\$29.95 (U.S. Funds).

NOTE - For prepaid subscriptions select
any two (2) reports as a bonus.

Circle No.04 on Reader Enquiry Card.

MICROCOMPUTERS and PHYSIOLOGICAL SIMULATION

James E. Randall, Indiana University
School of Medicine, Bloomington

Foreword by Arthur C. Guxton, University of
Mississippi School of Medicine, Jackson

"Here is where Dr. Randall's book is invaluable:
the background information on microprocessors,
combined with specific examples of biological
data simulated with various hardware and
software configurations, should allow any life-
science experimenter to progress rapidly from
the initial idea to a working simulation model."

BYTE

Contents:

Introduction. Microcomputer Components.
Operating Systems and Programming Languages.
Hardware Enhancements for Simulation. Repre-
sentative Microcomputers. Compartmental
Kinetics: A First Example. The Glucose Tolerance
Test. Cardiovascular System Mechanics. Arterial
Pulse Pressure. Vectorcardiography and the Limb
Leads. Distortion of Waveforms. Axon Action
Potentials. Cardiac Action Potentials. Formatting
Student Exercises. Index.

1980. 250 pp., illus. Paper 06128 \$14.50

Brochure Available from Publisher.

Price is subject to change without notice.
Price higher outside U.S.A.



Addison-Wesley

Advanced Book Program,
World Science Division
Reading, Mass. 01867, U.S.A.

London, Amsterdam, Don Mills, Ontario, Sydney, Tokyo

Circle No.11 on Reader Enquiry Card.

Some worthy publicity for elastin

A. Serafini-Fracassini

Biology and Pathology of Elastic Tissues.
Frontiers of Matrix Biology, Vol.8. Edited
by A.M. Robert and L. Robert. Pp.230.
(S. Karger: 1980.) \$74.25, SwFr.124.

THIS book attempts to expound, for the
benefit of graduate students and profes-
sional investigators, various areas of
scientific research which are contained
within the subject of molecular biology and
pathology of elastin. This is undoubtedly a
topic of considerable importance to bio-
chemical and medical workers, and one
which has attracted a large amount of
research over the past two decades. As
unambiguous interpretations of functional
properties of the protein still prove elusive,
it is of some value to have, brought
together in one volume, a collection of
critical evaluations of some recent
findings. The earlier contributions in the
book deal with structural and con-
formational aspects of the protein, while
the others discuss its implication in the
inflammatory response and in some tissue
disorders. The role of elastases in these
pathological conditions is adequately
outlined. However, in this context, one
would like to have seen included an up-to-
date account of the biosynthesis of elastin,
with particular reference to the nature of
the soluble biosynthetic intermediate and

to the biogenesis of the cross-links. Such a
chapter would have been of great value to
an outsider to the subject in appreciating
the significance of some structural features
of the protein and in understanding the
pathogenesis of specific disorders. Other
aspects which with advantage could have
been covered more fully are the physico-
chemical properties of elastin, and the
thermodynamics of its deformation to
balance Dr Urry's excellent contribution
on synthetic repeat polypeptides.

It is difficult to make a unified
judgement on this collection of papers
because the style of the contributions is so
varied. Some offer an accurate review of
the relevant scientific research, while
others mainly outline the results obtained
in the authors' laboratories, occasionally
with the exclusion of some lines of research
that might have merited discussion. This is
also reflected by the variable number of
references cited at the end of each chapter.

The book is well produced, adequately
illustrated and as a whole very readable.
Elastin has not yet enjoyed its share of the
general escalation in scientific pub-
lications, and I consider this a worthy
collection, one which could be read with
profit in its entirety by non-specialists. □

A. Serafini-Fracassini is Professor of Bio-
chemistry at the University of St Andrews.

Compleat diquat and paraquat

R.J. Hance

The Bipyridinium Herbicides. By L.A.
Summers. Pp.449. (Academic: 1980.) £33,
\$79.50.

ONLY two bipyridinium herbicides, diquat
and paraquat, are in widespread use, so it is
interesting to speculate why they have
stimulated a book of this size when larger
groups of herbicides have not. Perhaps
there are two reasons. The first is the
enthusiasm of Dr Summers for his subject.
The second is that, in comparison with
other herbicides, the peripheral synthetic
chemistry of these compounds is
challenging and their toxicology unusual.
Indeed, these topics occupy almost half the
book. The remainder covers the chemistry
of the respective diquatery salts,
structure-activity relationships, environ-
mental fate, occurrence and analysis of
residues, herbicidal use and modes of
action.

With 2,500 references cited in little over
400 pages the text is hard going in places
but, by intelligent use of tables, the author
has done his best to avoid producing a
catalogue. The only reasonable criticism
that might be made is that not all readers
will follow the logic underlying the
sequence in which subjects are presented.

For instance, the chapters dealing with fate
in biological systems and with residues are
separated from that on toxicology and
environmental effects which in many
respects is closely related. The layout
within chapters also occasionally seems
strange. There is a good example on p.394,
where a paragraph describing the most
effective ways of treating paraquat
poisoning is placed in the middle of a
section rather than at the end where it
would have more impact. However, the
book is sensibly indexed so the reader
should have little difficulty in finding
specific information.

Nowadays, books including a range of
specialist subjects tend to be multi-
authored. For one man to tackle such a task
successfully is an impressive feat of
scholarship and Dr Summers deserves
warm congratulation. His book
undoubtedly will become a standard
reference. It shows, incidentally, that there
remains a place for books of this type
alongside the computer-based information
retrieval systems that seem to threaten their
viability. □

R. J. Hance is leader of the Herbicide Group at
the Agricultural Research Council Weed
Research Organization, Yarnton, near Oxford.

26 March 1981

Reagan's non-existent science advisors

Why is the Reagan Administration dragging its feet on the appointment of a Science Advisor in the White House? More than two months have now passed since Dr Frank Press packed his bags and left. From time to time, this or that distinguished person has been mentioned as a possible successor, only to sink beneath the waves of gossip again. The President's budget request dutifully includes provision for a Science Advisor and a small staff, but at the present rate the new fiscal year will have begun before the Administration will be writing cheques against this line item. It is, however, inconceivable that the motive for the delay in appointing a successor to Dr Press is simply parsimony. Even the new stringent budget includes many more extravagances than this modest office. Nor is it likely that the failure to appoint a Science Advisor stems from a lack of qualified candidates. Although some of those mentioned in connection with this office are probably too accustomed to speaking out in public to suit the White House, there can be no enduring shortage of people able to take on the job. Moreover, it is unlikely that President Reagan's White House is fearful that anyone appointed to the post would have to walk in Dr Press's shadow for too long. By common consent, Dr Press did the job superbly, but the new Administration seems not to have been shy of appointing less well known people to other posts previously occupied by relative giants.

The most charitable explanation is that the new Administration has sought to avoid the encumbrance of a Science Advisor while putting together its budget request. And it is easy enough to understand what a nuisance such a person would have been in the past few weeks. The proposal to cancel the United States half of the joint mission, with the European Space Agency, to put a pair of satellites in polar orbits about the Sun would have evoked a string of memoranda explaining that such a course of action might permanently damage the chances of international collaboration on technical projects. Similarly, the Administration may have calculated, the plan to take the National Science Foundation out of science education would certainly have produced a reflex reaction from any Science Advisor anxious to keep his links intact with former colleagues in the wider world. More memoranda would have filtered over from the Old Executive Office. President Reagan's spirits must have sunk quite low at the prospect that a newly appointed Science Advisor, anxious to make his mark alongside the Killians and Wiesners, would have chosen to offer unsolicited and largely unusable (let alone unwelcome) advice about the new Administration's armoury and its lack of policy (so far) on arms control. Why not wait until the summer, when some of the dust will have settled, before inviting such unnecessary problems? On this charitable calculation, the next President's Science Advisor will be tapped on the shoulder some time about midsummer, when his or her plans to spend a vacation at Woods Hole, Aspen or in Europe are too far advanced to be cancelled easily, will take up office in the middle of the next frenzied budget process, and will not start writing awkward memoranda until fiscal 1983.

dominates everybody's thoughts. That is the charitable calculation of the new Administration's most charitable motivation.

Unfortunately, more sombre, even sinister calculations are possible. Is there, for example, a chance that the Administration has it in mind never to appoint a successor to Dr Press, not because it seeks to make a modest saving in the budget but because, as a matter of principle, it has decided that it does not want to be encumbered with such a person? Purists will quickly say that no modern administration would harbour such a thought. Has it not been amply demonstrated, ever since the end of the Second World War, that every administration needs a Science Advisor? Did not President Nixon's abolition of the President's Science Advisory Committee (not yet restored, even though Congress would like it back) show what harm can come to those who stab the research establishment in the back? One day, the calculation goes, there will be a constitutional amendment making the office of Science Advisor a part of the machinery of government, and from then on there will be no hesitation. Unfortunately, this simple calculation is wrong. It is probable that President Reagan has persuaded himself that if he cannot find a Science Advisor with all the virtues (distinction, discretion and an unwillingness to write memoranda unless asked), he will do without one. In that, he will be mistaken.

The case for having some kind of Science Advisor is subtle, administrative rather than an issue of principle. It is fashionable to say that science is so important that any president must have his own captive source, but that too stems from a mistaken view of what the world is like. The White House is not short of technical advice. Every government agency in Washington can supply a stream of memoranda to match whatever Dr Press's putative successor writes. The particular function of a Science Advisor in the White House is to hold the ring between the several warring agencies of government that ask contradictory (and impossible) tasks of every government — make unemployment go away, preferably by the encouragement of small business, make universities the powerhouses of innovation but less of a charge on the public purse, ensure peace by some technical device, but provide that if there is trouble in Central Europe, the United States will not be on the losing side. (A Science Advisor worth his salt will also be ready with opinions on whether the MX missile should be based in Utah or at sea, but may not again be required to advise, as in President Kennedy's time, on the best way to grow cranberries in Massachusetts.) In the last resort, however, what matters about the function of the Science Advisor in the White House is not the content nor even the quality of the advice rendered, but the fact of his or her existence. He or she is a lightning-rod, able to concentrate and then harmlessly to discharge the discontents of a powerful constituency. The constituency is not, however, all that powerful, these days, but short of funds and also of friends. An Administration with its wits about it would appoint a Science Advisor now, while it can still choose. If Mr Reagan delays much longer, he will find that he has less choice.

How not to run telecommunications

The British government seems well on the way to making a hash of its policy on telecommunications. Almost since last July, when the Secretary of State for Industry, Sir Keith Joseph, first spelled out the terms on which the old Post office monopoly would be

relaxed, the House of Commons has been struggling with a nasty compromise of a bill for the reorganization of the British Post Office. The proposal to separate the businesses of handling the mails and of managing telecommunications is sensible, if long

overdue. Most of the other provisions of the bill, especially those concerned with the new telecommunications business (trendily known as British Telecom) are unsatisfactory (see *Nature* 288, 424, 1980). Although British Telecom will no longer (if the bill becomes law) be able to decide for itself how to exercise all of its monopoly powers, these questions will be decided not by the balance of competition in a free market but by the Secretary of State and his civil servants, often with the connivance of British Telecom itself. Thus even on the relatively minor, or at least technical, question of what equipment telephone users may attach to their telephone circuits, the government has now agreed that recommendations by the British Standards Institution should be scrutinized (and no doubt quarrelled with) by British Telecom before being decided by the Department of Industry. That British telephone users will suffer from this fainthearted fudge is clear. Those principally concerned have so far failed to recognize that British Telecom will also be damaged.

Signs of trouble have already surfaced. Last week (19 March), the chairman of British Telecom, Sir George Jefferson, was complaining to an audience of equipment manufacturers that development is being hampered by the lack of capital. The complaint is all too familiar. Even in prosperous times, British governments have been unwilling to let the most buoyant of the nationalized industries modernize its network as quickly as technology would allow. When, as now, the cornerstone of financial policy is to keep a tight rein on public expenditure and thus on the need to borrow from the money markets, and when such funds as can be borrowed are increasingly being used to support industries (private as well as public) in danger of collapse, the pressure on British Telecom is greater than ever. The irony is that the investment even of publicly borrowed funds in the telecommunications network would be profitable, which is more than can be claimed for the other investments of borrowed money which the British government has in the past few months felt bound to make in British Steel, British Shipbuilders and British Leyland. As luck will have it, the continuing restrictions on British Telecom's freedom to borrow capital have now been sharpened because its own revenues have risen less quickly than expected, chiefly as a consequence of the recession. The result will be that British Telecom will have to cut back its already modest plans for development in the next two years. Some years from now, when the recession is eventually ended, Britain will be less well endowed with telecommunications than it could have been. The effect of government policy is thus to starve the future so as to feed the past.

But is not this impediment to the proper development of a communications network merely an unfortunate but unavoidable consequence of the recession? That is the simple response to Sir George Jefferson's plea that British Telecom should be liberated from the shackles that now bind it. The calculation is misleading. Can it ever be wise for a government playing entrepreneur to make unprofitable investments while refusing to invest in profitable projects? For private investors, that would be a path to bankruptcy. For governments, good sense is habitually clouded by political considerations. The British government has been investing in steelmaking, shipbuilding and the manufacture of motor cars in part because it cannot bring itself to face the social upheaval that the collapse of these traditional industries would cause in the parts of the United Kingdom in which they are concentrated. Unemployment, the government seems to have decided, is high enough already. So, too, it is. But to fail to invest in potentially productive enterprises such as telecommunications will entail that, when the recession eventually drags to an end, high unemployment will persist for longer than would otherwise be necessary.

How, in these circumstances, can the British government arrange for the development of the telecommunications network at the pace that technology will allow? For the government, the question does not come out of the blue. Sir Horace Barlow, Sir George Jefferson's predecessor at the Post Office, resigned a year ago because the Department of Industry would not agree to his proposal that the development of the telecommunications

network should be financed by allowing the Post Office to borrow some of its capital requirements from private sources. The Treasury with a simple logic insists that any debt that is ultimately guaranteed by the public purse must be counted as public borrowing. But there are ways round this over-simple rule, some of them already embodied in the Post Office Bill now before the House of Commons. For one thing, to the extent that the shading of the public monopoly will make it possible for the first time (in Britain) for terminal equipment to be purchased by private users and not purchased and leased to users by the monopoly, there is scope for reducing the public cost of telecommunications development. But the bill also includes the delphic provision that British Telecom will be allowed to borrow money from private sources for joint enterprises with private industry in which it does not have a controlling interest. In principle, these arrangements would allow expensive capital projects, the use of earth satellites within the telecommunications network for example, or the development of novel kinds of trunk transmissions routes, to be financed privately and then leased back to British Telecom. Precisely such a device is being considered by the British government as a means of allowing the British railways to be electrified without undermining forecasts of public borrowing in the years ahead. So why, while complaining at the chains that bind it, does not British Telecom point to the ways in which its own development could lighten the government's need to borrow money?

The explanation is straightforward but not creditable. Neither British Telecom nor the Department of Industry appears as yet to have adjusted to the opportunities provided by the shading of the public monopoly on telecommunications. While rightly urging that the pace of development in telecommunications should be faster — much faster — British Telecom seems not to have learned that it cannot do everything itself. If, however, its interest is the rapid development of a modern telecommunications network, it should welcome the chance of having as much as possible of the financial burden carried on other shoulders. It should, for example, welcome the prospect that it may no longer be required to buy terminal equipment from manufacturers simply so as to lease it on to users, tying up capital in the process. In practice, British Telecom has taken a much more defensive line, and is clearly planning to compete as fiercely as it can with such of the suppliers of terminal equipment as the Department of Industry may license in the fullness of time. Moreover, as the bill is now drafted, there is nothing to ensure the new monopoly will compete with upstart suppliers on terms that are objectively fair. Indeed, there is a danger that British Telecom will try to beat off the new suppliers to the network by devoting too much of its own resources to its first essay in competition, thus further hampering its capacity to invest in long-term developments. In the circumstances, British Telecom's pleas for more freedom to borrow money sound more hollow than intended.

The solution must lie with the government, the only participant in the reorganization of the British telecommunications network which has an interest that the old monopoly should be eroded and also that the development of the network should be less dependent on public borrowing. Unfortunately, the Department of Industry cannot ensure that British Telecom behaves sensibly while it has assumed for itself the role of final arbiter on trivial technical questions such as the specification and design of the terminal equipment that may be attached to telephone circuits. What is needed, even at this late stage, is that the British government should step back from the day to day regulation of the telecommunications network, delegating that responsibility to a less partial body able to take a judicial view of what constitutes fair competition between British Telecom and the rest of the telecommunications industry. Such devices have been found necessary elsewhere, in the United States for example, in striking a balance between the public interest and those of public monopolies. For the British government, and for British Telecom, the prize would be some relief from the problem of financing a technological development of great promise in the middle of a recession.

DNA guidelines for further attenuation

NIH to consider new voluntary provisions

Washington

Two prominent members of the National Institutes of Health (NIH) Recombinant DNA Advisory Committee (RAC) are suggesting a major transformation of the guidelines covering recombinant DNA research which would not only significantly reduce containment levels for a variety of experiments, but also convert the guidelines into a non-regulatory "code of standard practice". Dr Alan Campbell of Stanford University and Dr David Baltimore of the Massachusetts Institute of Technology claim that revoking the mandatory aspects of the guidelines is compatible with the experience gained since they were introduced five years ago.

The recommendation from Dr Campbell and Dr Baltimore will be discussed at next month's meeting of RAC. Also under consideration will be some less radical proposals suggested at a recent meeting of the chairpersons of local biosafety committees (*Nature* 288, 529, 1980), focusing primarily on reducing the paperwork involved in registering experiments conducted at the minimal P1 containment level, which is expected to receive the committee's support.

In a letter to Dr Donald Fredrickson, director of NIH, the two authors claim that since 1976, when the guidelines were introduced, neither experimental evidence nor theoretical arguments have been put forward to support the argument that recombinant DNA research poses any danger to human health. According to their proposal, the list of prohibited and exempt experiments would be retained, as would RAC. But all mechanisms which have been set up to enforce the guidelines as regulations — including, presumably, local institutional biosafety committees — would no longer be required.

Dr Campbell and Dr Baltimore argue that singling out some classes of experiments as needing higher levels of physical containment has been made at the cost of discouraging variety and innovation and thereby limiting access to useful knowledge. In contrast, they say, "the benefit is likely to be zero".

If NIH agrees to move in this direction, it is likely to proceed cautiously. At the last RAC meeting in January, the committee approved by nine votes to eight, with three abstentions, a proposal from Dr Winston Brill of the University of Wisconsin that the guidelines be changed to permit

recombinant DNA experiments using nonpathogenic prokaryotes and lower eukaryotes to be conducted at P1 containment.

However, in a notice published recently in the *Federal Register*, Dr Fredrickson says that in the light of the closeness of the vote, he is accepting in preference a compromise proposed by another committee member, Dr Susan Gottesman, which would allow the Office of Recombinant DNA Activities to lower containment levels for such experiments on a case-by-case basis, but without requiring RAC review.

One effect of the proposal of Dr Campbell and Dr Baltimore would be to throw into greater prominence suggested safety guidelines for microbiological and biomedical laboratories prepared by the Center for Disease Control (CDC) — which, like NIH, is part of the Public Health Service. However CDC itself has become a recent focus of controversy, following its publication last autumn of an updated draft of what it calls a "national code of practice".

The document, which has been widely circulated to universities and other research laboratories, is divided into two principal sections, and it is the second

section which has generated the most detailed criticism. This contains proposals for treating several types of organisms in a way which many microbiologists feel would be overly restrictive — in particular using viruses such as SV40. As now written, the guidelines "would put most virologists out of business" says Dr Harlyn Halvorson, chairman of the public and scientific affairs board of the American Society for Microbiology (ASM), which has been coordinating the views of several professional societies and is shortly to issue a detailed critique of the proposals.

CDC officials maintain that they are anxious to achieve a consensus on the guidelines, and plan a thorough revision in the light of the many comments that they receive. They also point out that the guidelines do not have the force of law outside public laboratories, and are intended more as a guide to local biosafety officers.

In reality, however, the CDC guidelines are likely to be widely adopted as a broad framework for dealing with potentially dangerous microorganisms, whether used in recombinant DNA experiments or not. In a letter to CDC director Dr William H. Foege, ASM asks CDC to broaden the basis for discussions with the biomedical research community — something that

Guidelines transgressor let off lightly

Washington

Dr Ian Kennedy, discovered last year to have been cloning Semliki Forest virus and not Sindbis virus, has been let off with a caution. The review committee set up by the National Institutes of Health (NIH) to investigate charges that he had infringed regulations covering recombinant DNA research in his laboratory at the University of California in San Diego has decided that no further action can be taken against him. Dr Kennedy is no longer conducting research funded by NIH and has no grant application under review for NIH funding.

The panel has also, however, recommended that, in the light of what it claims was Dr Kennedy's failure to comply with the NIH guidelines covering recombinant DNA research, during the next two years "his use of such techniques may be made subject to such additional conditions as the awarding [NIH bureau, institute or division] may require".

The NIH review committee throws little light on whether, when Dr Kennedy performed experiments using Semliki Forest virus which were at the time proscribed by the NIH rules, he did so intentionally or merely as the result of cross-infection. "It would not be fruitful for NIH staff to engage in an extensive examination of Dr Kennedy's laboratory notebooks or to interview his co-workers in order to attempt a resolution of complicated and disputed chronology."

The committee, nevertheless, faults Dr Kennedy for not seeking prior approval for recombinant DNA experiments in mouse L cells, from which transformed RNA was analysed, and says that Dr Kennedy failed to act responsibly.

Commenting on the report, Dr Donald Fredrickson, director of NIH, said that he accepted the committee's recommendations, and has asked both NIH officials and the university chancellor, Dr Richard Atkinson, to take appropriate action.

The review committee's report comes two weeks after Dr Martin Cline agreed to step down permanently as head of the division of haematology and oncology at the University of California, Los Angeles, following the university's discovery that he had used recombinant DNA molecules in treating patients with thalassaemia without the consent of the university committee responsible for protecting human subjects during clinical experimentation.

A statement issued by the university said that although both the patients and their families were told of the benefits and risks of the experimental treatment — which was carried out on one patient in Italy and another in Israel — neither university nor hospital officials were informed. The university's chancellor, Charles E. Young, added, however, that no harm had come to the patients treated by Dr Cline, whom he hoped would remain at the university.

David Dickson

CDC officials claim they would do

The question that remains is how this would affect the future of RAC. One option might be to open the committee up into a general "biosafety advisory committee", a proposal already supported by several committee members. Another would be to broaden its scope to consider some of the ethical issues raised by the use of genetic engineering techniques in medical experimentation — a direction which Dr Fredrickson is said to favour. Next month's meeting will provide some clues.

David Dickson

UK research councils

Spending too soon

Sir Geoffrey Allen, chairman of the Science Research Council, was let off lightly by the House of Commons Public Accounts Committee last week. The committee had hauled him in to explain why the council had overspent in the year 1979–80 by £4.1 million. Mr Joel Barnett, chairman of the committee, reprimanded Sir Geoffrey for errors of judgement, but seemed to accept his explanation.

The trouble began, said Sir Geoffrey, with the council underspending its budget



No Tower for Allen

by £2 million in each of the previous two years and with the promise in late 1978 from Mrs Shirley Williams, the Secretary of State for Education and Science, of an extra £33 million for the years 1979–80 to 1982–83. To make sure that unspent money would not be taken back, Sir Geoffrey decided to spend the extra allocation for 1979–80 quickly and asked universities to apply for research and equipment grants. But by July 1979 it was beginning to be apparent that the council might be overcommitting funds, especially as the new Conservative government, elected the previous April, had said that it would not honour Mrs Williams's promise in full.

By September things had got worse. The council then knew that it would only receive £5 million of the £7.5 million it had been expecting for 1979–80. An attempt to withdraw some of the extra grants awarded

was unsuccessful because equipment had already been delivered and money spent. By October, the council was forecasting excess expenditure of £7 million by the end of the financial year. In the event, it managed to save £2.9 million, mainly on international subscriptions.

The council had been caught out because equipment manufacturers were able to deliver on time and because universities, with cash flow problems of their own, submitted bills promptly. Work on the spallation neutron source at the Daresbury laboratory had also been completed earlier than expected — the council had been hoping that the bills would fall in the succeeding year.

Mr Barnett suggested that Sir Geoffrey's stratagem to ease cash flow by postponing payments amounted to deception of Parliament. But officials from the Department of Education and Science and the Treasury were more sympathetic in their evidence. The problem is that the council, like other government-funded bodies, has to plan its annual expenditure within a very small margin of error. Sir Geoffrey's request that the error be raised from less than 0.2 per cent to 1 per cent, however, was quashed by the Treasury witness who said that an announcement had been made the previous day that no easing of cost margins was to be granted for any public agency.

The council's overspending for last year is bound to make this year's financial problems even more delicate, especially with the Public Accounts Committee breathing down the council's back. In the meantime, Sir Geoffrey expects to be issued with a formal reprimand. He is unlikely to be sent to the Tower of London.

Judy Redfearn

Polish academics

Degrees in question

Edward Gierek, former First Secretary of the Polish United Workers' Party, made a comeback to public life recently — fighting to defend his right to an MSc. In an interview published in the Katowice Party daily, *Trybuna Robotnicza*, Mr Gierek complained that he had been subject to a prolonged campaign of denigration.

Mr Gierek was referring to allegations made by the Krakow Party daily *Gazeta Krakowska*, that the Krakow Akademia Gorniczo-Hutnicza (Mining and Metallurgy Academy — AGH) had improperly awarded scientific degrees to prominent political figures including Zdzislaw Grudzien (former Katowice Party chief), Franciszek Szlachcic and Stanislaw Kowalczyk (former Ministers of the Interior) and to Mr Gierek himself.

Meanwhile, the Solidarity chapter at AGH had published a list of people alleged to have gained degrees improperly, claiming that in the early post-war period special crammer courses were introduced in a number of technical specialities — but

that in practice only Party and government functionaries were able to take advantage of them.

Mr Gierek says that he was only one of several thousand people who had taken advantage of the courses intended to allow people with long industrial experience to get a degree in engineering "without regular studies".

The rector of AGH is Dr Roman Ney, who last autumn became a candidate member of the Party politburo. He was swift to respond to the suggestion of irregularities, and set up a special commission to investigate the allegations. In Gierek's case it found no evidence of malpractice. AGH had on file a copy of his diploma as a mining engineer, specializing in the exploitation of coal deposits, and duly approved by the examination board. Regarding the other allegations that in 1957–61 Grudzien, Szlachcic, Kowalczyk and others had been allowed to complete four-year extramural courses in the economics of metallurgy in just over one year, the AGH department of metallurgy simply reported that these particular courses had been established on the orders of the Minister of Education.

Further queries have been raised about the academic status of Andrzej Jaroszewicz (son of the former Prime Minister). On 9 March, Krakow Technical University set up a special commission to investigate the circumstances surrounding his external degree in engineering.

These swift responses to hints of irregularity are in sharp contrast to the attitudes of officialdom which led, in 1976, to the resignation of the then President of the Polish Academy of Sciences, Professor Janusz Groszkowski, from all his public offices. Professor Groszkowski had attempted to intervene in the case of Mrs Aleksandra Hankus, who in 1964 had complained to the Krakow Technical University that a Mr Tibor Petrys had stolen her scientific results. But the Technical University refused to take any action against Mr Petrys, granted him his DSc and later helped him prepare evidence for a libel suit against Mrs Hankus.

After an 11-year struggle — during which she served an 11-month sentence for libel and was sent by the judiciary for psychiatric tests — the Supreme Court finally vindicated Mrs Hankus, stating that she had proved that Petrys had "usurped her scientific work and cheated in obtaining the degree of Doctor of Sciences". When Professor Groszkowski tried to bring a legal action against Mr Petrys, the Procurator General blocked the case for a year, since, according to Professor Groszkowski, Mr Petrys still enjoyed the support of "some organs of justice". But eventually Mr Petrys' degree was revoked.

In September 1980 Professor Groszkowski at last re-entered public life with an open letter to the academy, calling for a return to true academic standards in Polish science.

Vera Rich

US universities

Commercial interest*Washington*

Legislation passed last year by the US Congress to encourage the scientific community to take a greater interest in the commercial exploitation of their research — for example by giving universities and small businesses the right to patents resulting from federally-sponsored research — is already bearing fruit. Spurred by the new patent laws, Stanford University is considering whether to invest university funds in companies set up by faculty members to exploit their research. And the National Institutes of Health (NIH) are expected to reverse a 20-year-old policy which excludes scientists who work for profit-making companies from applying for competitive research grants.

At the same time, there are renewed pressures in Congress for legislation to offer private corporations tax incentives for supporting basic research in universities. One attraction for the universities stems from the increasing constraints on federal funding, but industrial support for university research — at present only about 3 per cent of the total — is unlikely to become a substitute for the federal government.

New NASA man*Washington*

The Reagan Administration is expected to announce shortly the appointment of Dr James Beggs, at present executive vice-president for aerospace at General Dynamics Corporation, to be the next administrator of the National Aeronautics and Space Administration (NASA). Dr Beggs was NASA's associate administrator for advanced research and technology in the late 1960's and before joining General Dynamics had acted as Under-Secretary in the Department of Energy. His name was informally mentioned by Vice-President George Bush during a visit to the Kennedy Space Centre last week, and a formal announcement about Dr Beggs' nomination is expected from the White House this week.

Meanwhile there is still no news about the post of science adviser and director of the Office of Science and Technology Policy (OSTP). A decision by the White House to include funds for OSTP in its budget request for 1982 indicates that no official decision has yet been made to eliminate the office or its activities. However, there are several rumours in Washington that the science adviser's post may be downgraded, and that OSTP could be absorbed into the Office of Management and Budget, with which it worked increasingly closely during the Carter Administration. **David Dickson**

The Stanford proposal was put to a meeting of the university's committee on research two weeks ago by Dr William Massey, professor of business administration and vice-president for business and finance. In particular, he spoke in favour of setting up a Stanford investment trust which would operate partly on capital provided by the university and partly on capital raised outside. Such a trust would make "modest" equity investments in the early financing of high-technology companies whose products derived in some way from work at Stanford.

Several scientists at the Stanford meeting, among them Dr Edward Garvin, an adjunct professor at the Stanford Linear Accelerator Center, were concerned about the potential impact on research activity of the new proposals, particularly if it meant that the potential ability to earn money for the university became a major factor in making research choices or faculty appointments.

At NIH, the change in rules covering applications for research grants is expected to be announced in the *Federal Register* within the next few weeks. It reverses a restriction applied to all funding by the Public Health Service in 1962, partly as an attempt to ensure that research funds were used to help build up university and medical schools, partly because of some well-publicized abuses.

A return to the situation of before 1962, which would not require ratification by Congress since the decision was essentially administrative, has been under discussion for several years. The document setting out the change in procedures is now said to be awaiting merely the signature of Health Secretary Mr Richard Schweiker, who is not expected to raise significant objections.

The legislation had some important administrative implications. For example, changes in the patent law now required NIH to keep track of all patents awarded to universities from federally-sponsored research, in order that the institutions should be able to exercise "march-in rights" if they consider that a patent is not being sufficiently exploited.

Dr Charles Lowe, head of the institutes' Office of Medical Applications of Research, warned that such changes could have important implications, and steps might be necessary, for example, to protect the integrity of the peer review process. His warnings were echoed by Dr Philip Siekevitz, professor of cell biology at Rockefeller University in New York, who said there was a danger that biomedical research priorities would be shifted from social need to return on investment.

Dr Donald Fredrickson, director of NIH, said he expected that many of the changes described by Dr Lowe would come about and that NIH would in particular be looking closely at the experiences of the National Science Foundation in forging ties between industry and universities.

David Dickson

UK astronomy

Change of direction

Responsibility for the construction of the La Palma observatory in the Canary Islands, Britain's hope for good optical astronomy in the Northern Hemisphere, will be handed over to Professor Alexander Boksenberg, now professor of physics at University College London, on 1 October. Professor Boksenberg will then succeed Professor F. Graham Smith as director of the Royal Greenwich Observatory, which is responsible for building and subsequently operating the La Palma site. Professor Smith, a radioastronomer, will succeed Sir Bernard Lovell as professor of radio astronomy at the University of Manchester and director of the Jodrell Bank radiotelescope.

The La Palma observatory, which will house four major telescopes, is likely to be the Greenwich observatory's largest project for a long time to come. The Science Research Council, which made the new appointment, hopes that one of Professor Boksenberg's first tasks will be to chivy the Spanish authorities, who have been responsible for some delay, into speedier action. But Professor Boksenberg sees his main task as encouraging astronomers in British universities to take the fullest possible part in the new observatory. He would like to see their involvement extend beyond the use of the telescopes for observation to the design of instruments and software, either in their own laboratories or at the Greenwich Herstmonceux site.

Since the Royal Greenwich Observatory took on the construction of the La Palma observatory, its own research effort has gradually diminished, a trend that Professor Boksenberg would like to see reversed. He hopes to encourage more astronomers to join his organization and, by collaboration with universities and Britain's other major observatory, the Royal Observatory Edinburgh, to transform the astronomy community into "a happy family".

Professor Boksenberg's previous experience seems particularly relevant to an organization involved in both instrument-building and astronomical research. His expertise in building image intensifying instruments has won him collaboration on several major worldwide telescopes. The image photon counting system, his big success, earned him observation time on the Kitt Peak and Palomar Mountain Observatories in the United States. He is at present building instruments for the faint object camera, Europe's contribution to the space telescope.

His observations have been mainly extragalactic, of quasars in particular. As well as using groundbased optical telescopes, he has also worked with data from and designed instruments for the

International Ultraviolet Explorer satellite Professor Boksenberg hopes to retain an interest in astronomy at University College London by becoming a visiting professor there **Judy Redfearn**

Oak tree wilt

Europe in dread

Brussels

The European Commission is about to launch a research programme to find a cure for the tree disease oak wilt. And a system of import controls is already in force to ensure that wood from the United States does not trigger off an epidemic as devastating as Dutch elm disease.

Experts on both sides of the Atlantic fear that if oak wilt were to reach Europe, it would in the European climate be the death knell of the oak tree. The fungus has spread slowly in the United States and is concentrated in the areas below the Great Lakes. Oak wilt is a fungus (*Ceratostoma fagacearum*) whose spores are thought to be transmitted either by the scullitus beetle or by root grafting (roots intertwining underground). The fungus clogs the tree's vessels, blocking the movement of the sap.

Controversy has surrounded the necessity for and nature of the import regime. Britain and Ireland have been in favour of tight restrictions on imported wood. But aggressive lobbying by the German furniture industry has led to relaxation for some continental industries. The German industry requires oak logs with the bark intact for veneer furniture, and the United Kingdom's Forestry Commission has stressed that the bark poses the greatest dangers of transmission.

As things are, the logs have to arrive accompanied by a certificate to show that the wood comes from uncontaminated areas of the United States. The logs can then enter the Community only through a limited number of ports and cannot then be re-exported to another member state.

Further precautions are nevertheless being taken. The Community's import regime requires that the wood is square-edged, has been dried to remove 20 per cent of the moisture and is subject to heat and air treatment. In case of doubt, the wood can also be ponded or submerged in water to kill remaining fungus. The problem for the furniture industry is that such processes do the wood no good.

The research programme is being mounted in collaboration with American and Canadian research institutes (although Canada is still unaffected). Five contracts are about to be let by the Commission to plant research institutes in the United Kingdom, Germany and Belgium. Details of the programme have not been released, but one contract will go to the Forestry Commission. The total cost of the programme is estimated at 570,000 European Units of Account (£0.54 = 1 EUA) with Europe contributing just over a half, the United States 200,000 EUA and

Canada 50,000 EUA.

Although oak wilt has been known for the past 40 years, work is still going on to map the afflicted areas. Efforts will also be devoted to gaining a greater understanding of how the disease spreads, present evidence suggests that oak wilt would spread far more quickly in Europe than in North America.

Methods exist to kill the fungus in the wood but the more important challenge is to find a way of killing the fungus in the living tree.

Problems have arisen in coordinating the timing of the research, as the Americans are in less of a hurry. European research is also handicapped by the lack of infected wood, for obvious reasons. Relations between Europe and the United States are also coloured by the knowledge that Europeans were responsible for exporting Dutch elm disease to North America in the first place.

Jasper Becker

Soviet higher education

Candidates change

The Soviet system of higher academic degrees is due for radical overhaul, according to an article in the prestigious *Literaturnaya Gazeta* last week. On the eve of the Annual General Meeting of the Academy of Sciences of the USSR, it published a major critique of the system under which the degree of Doctor of Science is now normally awarded to mature and well-known scientists approaching their fiftieth birthdays.

The author of the article, L. Orlov, is a typical product of the system. He heads a research laboratory employing 50 scientists, and has published more than 80 papers. As a senior scientific worker, holding the degree of Candidate of Science, he earns 300 rubles per month. To obtain the extra 100 rubles a month which go with the title of Doctor of Science, he is now having to rework his research results into a formal thesis, to be defended in exactly the same manner as he defended his Candidate's thesis years ago.

Orlov does not deny that the two-stage degree system had its virtues when it was originally introduced 50 years ago, when the rapid expansion of Soviet science and the lack of trained personnel meant that young scientists who had not yet completed their doctoral research were employed in posts which would normally have gone to Doctors of Science. Under such circumstances, some pre-screening of these "doctoral candidates" was necessary.

Orlov considers the double-examination system no longer relevant since Soviet science and higher education has now reached an extremely high standard. He suggests that the two-stage system is actively harmful to science. Not only do the examiners sometimes let through a young scientist whose thesis is not entirely satisfactory, "since it is, after all, only a

Candidate's thesis", but the need to produce two such pieces of original research in one academic lifetime can seriously hamper a scientist's creativity.

The current Five Year Plan, approved by the Party Congress last month, specifically calls for flexibility in research and development in order to meet the demands of the scientific and technological revolution, and for improvements in training of scientists and workers in higher education. A provocative article in a leading journal is a traditional Soviet way of introducing a possible reform to public opinion, and is usually followed by a spirited "press debate". Orlov's proposals will probably attract considerable sympathy among scientists and academics. Not only would the adoption of a single higher degree eliminate the need to write a second thesis, it would also eliminate the chagrin arising from the present system, which as Orlov notes, gives formal parity between a Soviet "Senior Scientific Worker" who at forty plus has gained his doctorate after 15 years hard work, and some "27-30-year-old young Doctor of Science from East or West Germany".

Vera Rich

Irish research and development

Marching to Georgia

Ireland is to set up a new contract research institution in the hope of persuading the 800-odd foreign companies based — tax-free — in the country to do some research and development there at last. The new institution, announced last week, and to be called the European Research Institute of Ireland (ERII), will be set up in Limerick with the help of America's Georgia Institute of Technology and will offer 50 per cent government grants to companies wishing to place contracts with it.

Ireland's Industrial Development Authority, which is responsible for the arrangements, foresees ERII as leading Ireland into its third stage of industrial development. The first was purely domestic, the second involved the establishment of foreign industry, and the third will involve research and development placing Ireland among the modern technical countries. The problems in reaching that third stage have been first, that most Irish researchers travel abroad to do their work, to America or Europe, and second, that whereas other countries offered tax incentives to companies to invest in research and development, Ireland could offer none as the firms paid no tax anyway.

The new institute is seen as the solution to these problems, with the grants substituting for tax relief, and high salaries to attract expatriate researchers back to Ireland. The Industrial Development Authority calculates that it will be paying staff at Limerick enough to match the

standard of living they might enjoy if they were based, say, at the Battelle Institute in Frankfurt.

ERII will start formally on 1 June, when two members of the Georgia Institute of Technology Engineering Experiment Station begin work in Limerick. Initially they will be housed in Ireland's National Institute for Higher Education, where president Dr Edward Walsh is said to be playing an "aggressive" part in forging links between university and industry. Recruitment is under way for eight engineers, scientists and computer specialists, but in three years' time, when ERII will move to its own buildings, it is planned that staffing levels should increase to 80.

To finance ERII, the Industrial Development Authority is paying for the buildings (£1.5 million); and the balance will come from such sponsors as Guinness Ireland, the Bank of Ireland, Allied Irish Banks and the industrial Rohan Group. Nevertheless the authority is seeking another £500,000 to launch the institute "effectively", and would like more companies to contribute to the consortium.

Robert Walgate

UK universities

Jobs to be cut?

One week after the British government's budget, the University Grants Committee has plucked up courage to make some guesses on the quantitative implications for universities. Dr Edward Parkes, the chairman, told the Public Accounts Committee of the House of Commons last week that to meet an 11 per cent cut in real terms in the universities' grant over the next three years, 3,000 academics and 4,000 other staff would have to lose their jobs. Redundancy payments to academics, said Dr Parkes, could be somewhere between £40,000 and £80,000, although realistic figures would not be known until cases had been through the courts. The total cost to the universities could be £100–200 million.

Dr Parkes has since said that in addition to 3,000 redundancies, 2,000 jobs would have to be lost through natural wastage. He estimates that 3,000 academics will leave their posts in the next three years, either through retirement or to take posts else-

where. His current plan allows for 1,000 vital posts to be refilled. A total loss of about 5,000 academic jobs represents just under 15 per cent of the academic workforce. The percentage of job losses will have to be greater than the percentage cut in the budget because economies in other areas of expenditure are virtually impossible, according to Dr Parkes.

The Association of University Teachers disagrees with that view. It is preparing a scheme to avoid large numbers of redundancies by relying on natural wastage, an early retirement option at 50 and spending one-third of the equipment grant on staff. But the scheme is unlikely to appeal to the grants committee which holds that freezing all posts would be the worst option. The problem is that those departments with the highest rate of turnover of staff — medicine, science and engineering — would suffer most, and in many cases, those are the very departments that the government would like to see least damaged. The committee prefers to maintain minimum flexibility in making new appointments and create the minimum number of redundancies.

The Science Research Council, which provides research grants to individual academics, is particularly concerned that good research should not disappear. Ever optimistic, it says that it may be able to support a few outstanding individuals who would otherwise have to go. Such support, however, is unlikely to be very significant.

The Committee of Vice-Chancellors and Principals has estimated that the grant for 1983–84 will be £80 million less than that for this year and that the loss from overseas students' fees could remove a further £70 million, a total sum short of conservative estimates of redundancy payments. The estimated costs may well have come as a surprise to the government, which is showing signs of reconsidering its plans. Sir James Hamilton, Permanent Secretary at the Department of Education and Science, told the Public Accounts Committee that his department was still attempting to work out with the University Grants Committee how the economies could be achieved. Although the budgets for 1981–82 and 1982–83 are fixed, he said, the figure quoted for 1983–84 is still only an estimate, suggesting that there may be room for further negotiation.

Universities, however, are unlikely to make up lost income by taking in more students. Contrary to the impression given in the white paper on public expenditure and by Mr Mark Carlisle, Secretary of State for Education, Sir James Hamilton told the Public Accounts Committee that the government will not increase the grant to universities that overstep their quota of students. The plan is that student numbers should remain roughly constant, implying greater competition for places as the number of eighteen-year-olds in the population continues to increase over the next few years.

Judy Redfearn

Bulgarian astronomers look up and abroad

Bulgaria's new National Astronomical Observatory on Mount Rozhen, opened on 13 March 1981 by President Todor Zhivkov, was several times referred to in the inaugural speeches as a "pre-Congress gift to Bulgarian science". The term is appropriate, for the theses for the forthcoming Congress of the Bulgarian Communist Party state that in the new Five-Year Plan "primary importance should be paid to fundamental research".

This is in sharp contrast to the policies of many other East European countries, which tend to give first priority to application. According to Kiril Serafimov, head of the Bulgarian Committee for Space Research and Exploitation, "we are also a European country and understand that without fundamental science nothing is possible".

Speaking at the opening ceremony, Dr Serafimov inevitably referred to Bulgaria's "thousand-year-old" interest in astronomy — the country is this year celebrating its 1,300th anniversary of statehood. The observatory has three telescopes and a large data-processing and computer unit. The main telescope has a 200-cm hyperbolic mirror with a Coude-Richey-Cretien focusing system and will be used primarily for work on binaries and variable stars. The second, a 70-cm Schmidt instrument (effective diameter 50 cm) with a 4° x 4° field of view, will be used for preliminary search and scanning work. The smallest, a 60-cm instrument, is especially adapted for photoelectric work and brightness determinations.

The observatory is the most southerly in the socialist bloc (and, according to Bulgarian claims, the second largest optical observatory in Europe). The Rozhen team

expects to cooperate closely with the Soviet observatories, especially the Crimean Astrophysical Laboratory and the Special Astrophysical Observatory in the Caucasus, on approximately the same latitude. Intra-Comecon cooperation is well represented at Rozhen — the research programme has been worked out in collaboration with teams from the Soviet Union, the German Democratic Republic and Poland. (The optical equipment has been supplied by Carl Zeiss (Jena).) There has also been some French participation in drawing up the programme, while the data-processing equipment is largely of Western origin.

Vera Rich



Mount Rozhen in snow

CORRESPONDENCE

Toxicity testing

SIR — The dramatic fall in Fisons' shares as a result of the withdrawal of Proxicromil emphasizes the high risks involved in drug development (*Nature* 12 March, p.81). Drug development programmes depend on animal tests for assessing long and short term toxic effects which are notoriously unreliable and expensive.

The effects of drugs on animals cannot be assumed to be the same in humans, irrespective of whether results indicate that a drug is "safe" or "harmful". Legislators should accept this fact and release toxicologists from the present "check-list" testing system which has replaced scientific initiative with standard tests of dubious relevance which are carried out on ever increasing numbers of animals.

On scientific, economic and humanitarian grounds current toxicity testing protocols should be refined to ensure the most efficient use of animals and to incorporate valid non-animal screening tests wherever possible.

LESLEY SARGEANT

*Fund for the Replacement of Animals in Medical Experiments,
London, UK*

Inherited tolerance?

SIR — I would like to suggest a more straightforward explanation of the fascinating findings of Gorczynski and Steele¹ than the one they have opted for. There is a strong possibility that mouse semen contains some cells other than sperm — lymphocytes, for example. In the mouse, semen is ejaculated intra-cervically, rather than intra-vaginally as in humans. Thus any non-sperm cells will be directly introduced into the uterus in the mouse. Not much is known about the fate of such cells although it appears that sperm are rapidly removed by phagocytosis. The pregnant uterus, however, is to some extent tolerant of foreign cells. Short and Yoshinaga², Wilson³ and Lions⁴ have described experiments which show that heterologous tumour cells can persist for several days in the rat uterus and that these can enter the mucosa only at the time of implantation in pregnancy. Thus the endometrium becomes receptive to foreign cells at the time of implantation, and so non-sperm cells present in semen could presumably become established in the endometrium and might then enter the embryo after implantation.

If we then suppose that tolerance involves clones of cells in a manner analogous to immunity and that some of the non-sperm cells are competent for the transmission of the tolerant state, the embryo might be rendered tolerant to the same antigens as the father. Moreover, if the chances of such an invasion of the embryo were less than 100 per cent, tolerance of a particular antigen would be found in only a proportion of the embryos and some would be tolerant to both of two antigens whilst others would be tolerant only to one. Since the postulated cells conferring tolerance are differentiated they would play no part in embryogenesis but would be carried and perhaps multiplied in the embryo. In subsequent generations, the chances of transmission of the cells would be finite and

the incidence of tolerance to specific antigens would decline. In the mothers, the return to the non-pregnant state after parturition would, presumably, involve the rejection of any foreign clones of cells in her tissues. Thus, subsequent matings to non-"tolerized" fathers would not result in the transmission of previous tolerant states.

In this way, the essential aspects of the Gorczynski and Steele experiments might be explained along rather less controversial lines than at present. At the very least, the possibilities raised here must be eliminated before we venture to compromise the doctrine of Weismann and begin resurrecting the long dead Lamarck.

K.W. JONES

*Department of Genetics,
University of Edinburgh,
Edinburgh, UK*

1. Gorczynski, R.M. & Steele, E.J. *Nature* **289**, 678–681 (1981).
2. Short, R.V. & Yoshinaga, K. *J. Reprod. Fert.* **14**, 287 (1967).
3. Wilson, I.B. *Proc. zool. Soc. Lond.* **141**, 137 (1963).
4. Lions, J. thesis, Univ. Cambridge (1970).
5. I thank Johnathan Aitken for discussion of these ideas and for the references cited.

Exhibit dismay

SIR — I have been following with great interest and some dismay the extended correspondence in *Nature* about the nature and significance of new exhibits in the British Museum (Natural History). I wish to comment on recent letters from three colleagues and fellow-countrymen of mine^{1–3}, one British colleague⁴ and one French colleague⁵.

McKenna¹ and Nelson² speak out against Halstead and in favour of cladistic or Hennigian systematics, to which a considerable part of the new dinosaur exhibit is devoted. Both McKenna and Nelson are now at the American Museum of Natural History, with which I was associated for many years. They belong to a small group now there which is representative neither of the staff of the museum as a whole nor of the majority of American palaeontologists, who, like me, reject Hennigian procedures.

Gould³ is at the Museum of Comparative Zoology, with which I was also associated for some time. He proclaims that he is not a cladist and "did not care for" the cladistic exhibits at Cromwell Road. He also claims that his theory of punctuated equilibrium is not a Marxist plot and notes that the co-author of the theory, Eldredge, is not a Marxist. I do not wish to discuss Marxism in this connection, but I do note that the punctuated equilibrium of Eldredge and Gould is only one facet, taken to an extreme, long since incorporated in the synthetic or if you like, neo-Darwinian, theory of evolution.

Howgate⁴ named me as being "by no means averse to the idea of leaps," that is, to saltatory evolution. He quotes as evidence a statement by me in a book published in 1949 (not 1976 as he has it) that rates of evolution in the horse family varied considerably. I did not state or imply that the rates were ever saltatory, that is, instantaneous.

Janvier⁵, of the Laboratoire de Paléontologie des Vertébrés of the University of Paris claims that cladism is "directly descended from Darwin's ideas on phylogeny and systematics." In context this implies that Darwin was a saltationist, which is not true.

He also notes that some French scientists (Grassé as an example) are anti-Darwinian. In fact French biologists, with the exception of a few radicals, have always been anti-Darwinian, at times for no clear reason other than chauvinism.

G.G. SIMPSON

*Department of Geosciences,
University of Arizona, Tucson, USA*

1. McKenna, M.C. *Nature* **289**, 626 (1981).
2. Nelson, G. *Nature* **289**, 627 (1981).
3. Gould, S.J. *Nature* **289**, 742 (1981).
4. Howgate, M. *Nature* **289**, 344 (1981).
5. Janvier, P. *Nature* **289**, 626 (1981).

Evolving theory

SIR — The letter from H.W. Ball and 21 colleagues (*Nature* 12 March, p.82) contains the following astounding statement: "But the theory of evolution would be abandoned tomorrow if a better theory appeared".

The lapse of time between the appearance of a "better theory" and the abandonment of its predecessor can be established with fair accuracy in many topics of science. The history of science provides a large number of such events. Accordingly, the lapse of time between the appearance of the new and the relinquishing of the old might be regarded as one of the many facts of science. Your 22 correspondents, apparently seriously, estimate the magnitude of this lapse of time as one day. One day! They declare their devotion to fact and their abhorrence of prejudice. Are their other facts of a similar colour I wonder, or does evolution evolve more rapidly than any other field of science?

R.V. HESKETH

*Berkeley Nuclear Laboratories,
Berkeley, Glos., UK*

SIR — Bertolt Brecht put these words into the mouth of Galileo¹:

"And what we discover today we shall wipe off the slate tomorrow and only write it up again once we have again discovered it. And whatever we wish to find we shall regard, once found, with particular mistrust. So we shall approach the observation of the sun with an irrevocable determination to establish that the earth does *not* move. Only when we have failed, have been utterly and hopelessly beaten and are licking our wounds in the profoundest depression, shall we start to ask if we weren't right after all, and the earth does go round."

As far as evolution is concerned, I think most biologists would agree that "the earth does go round", there being no better way of accounting for the facts *at present*. So how are we to understand the over-cautious round robin from the British Museum? There is a world of difference between "If the theory of evolution is correct . . .", and ". . . evolution is a fact, proven to the limits of scientific rigour." Most of us, like yourself, would wish to be more positive than the first but less dogmatic than the second. In these days of proliferating crazies, the British Museum should not risk giving them more ammunition.

P.L. MILLER

Queen's College, Oxford, UK

1. Brecht, B. *Life of Galileo* (translated by J. Willett) (eds Willett, J. & Manheim, R.) (The Collected Plays Vol.5, pt 1, Eyre Methuen, London, 1980).

NEWS AND VIEWS

Genes pieced together — exons delineate homologous structures of diverged lysozymes

from Peter J. Artymiuk, Colin C.F. Blake and Albrecht E. Sippel

GILBERT¹ has proposed that the arrangement of eukaryotic genes in a mosaic of introns and exons might facilitate the evolution of new proteins. He suggests, for example, that if exons correspond to functions, recombinations within introns could put together new proteins with new combinations of functions. Blake² has pointed out that if exons also correspond to folded protein structures the new protein is more likely to achieve structural stability by being a combination of folded sub-structures.

If these ideas are correct, we would expect to discover in different proteins folded units of common structure and function, derived from common exons, combined with units of different structure and function. Examples may already be to hand in the dehydrogenases³, and in triose phosphate isomerase, pyruvate kinase⁴ and KDPG-aldolase⁵, but in the absence of the details of their exon/intron patterns this cannot be confirmed. The definition of four exonic regions in the hen lysozyme molecule⁶ allows a comparison of these regions with the regions of structural and functional homology in T4 phage lysozyme^{7,8,10}. This comparison is, intriguingly, consistent with the hypotheses of Blake and Gilbert, although as an isolated first attempt it cannot be adduced as proof.

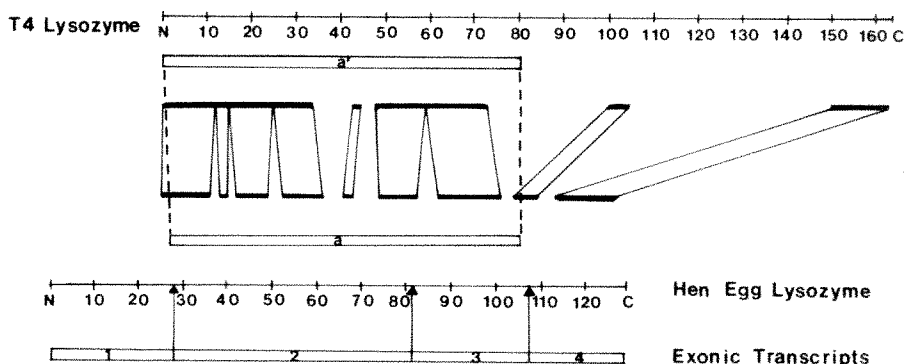
Determination⁹ of the structure of T4 phage lysozyme did not at first suggest that there was a close resemblance to hen egg-white lysozyme. Rossmann and Argos⁷ later showed that substantial portions of

the molecules were structurally similar. A number of small deletions in one molecule (or insertions in the other) made residues 1–73 of T4 lysozyme overlap with residues 25–101 of hen egg-white lysozyme. Two large insertions in the T4 chain brought out homologies in residues 100–105 and 150–163 of T4 with residues 104–109 and 113–126 of hen. The structural equivalence is shown schematically in the figure. In all, 78 α -carbons were shown to be equivalent with a r.m.s. difference of 4.1 Å. Using somewhat different techniques that do not allow insertions or deletions in the chains, Remington and Matthews⁸ confirmed the similarity and showed that residues 27–109 could be made to overlap with residues 1–80 of T4 lysozyme with a r.m.s. α -carbon difference of 6.0 Å. It is important to note that the structurally equivalent regions correspond to, at most, two-thirds of the hen lysozyme molecule and less than half the T4 lysozyme molecule, and for each lysozyme large contiguous regions of chain have no equivalent in the other molecule.

Rossmann and Argos⁷ noted that the equivalent regions encompass the active site regions of each of the molecules and showed that the binding sites for an *N*-acetylglucosamine monomer were only 0.8 Å apart in the two enzymes. Recently Matthews and his colleagues (see this issue of *Nature* p.334) have extended this

positional similarity to a bound tri-*N*-acetylglucosamine moiety, and to Glu 11 and Asp 20 of T4 lysozyme with the catalytic Glu 35 and Asp 52 of hen lysozyme¹¹. Furthermore, they suggest that sugar C is bound in a similar mode in the two enzymes and that sugar D is distorted in T4 lysozyme by a close approach of its 6-hydroxyl to protein groups similarly positioned with respect to the binding groups of sugar C, as in hen lysozyme¹¹. These observations imply that the structural similarity is accompanied by a similarity in substrate binding and catalytic mechanism. Although each separate piece of evidence is circumstantial, their combination argues strongly for an evolutionary relationship, or more unlikely, an extreme case of convergence in structure and function between the two lysozymes.

Sippel and his colleagues⁶ have recently determined the exon pattern in the hen lysozyme gene. They have found that the gene contains four exons, corresponding to residues –18 to 28 (the signal peptide of prelysozyme and residues 1–28 of mature lysozyme), 28–81/82, 81/82–108 and 108–129. They note that although the exon pattern is related to some degree to the structural division of the hen lysozyme molecule, its relation to the functional units of the enzyme is better established. For example, exon 2, which codes for residues 28–81/82, includes the catalytic residues and a cluster of amino acids that bind sugars C, D, E and F of the oligosaccharide substrate and could



A schematic diagram showing the relation between hen egg lysozyme and T4 phage lysozyme. The solid bars in the centre show the residues which are considered equivalent by Rossmann and Argos⁷. The rectangles marked a and a' define the 80-residue segments aligned by Remington and Matthews⁸. The lower rectangles and arrows mark the exonic transcript regions defined by Jung *et al.*⁶.

perhaps function as a primitive glycosidase. Exon 3, which codes for residues 81/82–108, corresponds to a protein fragment that, combined with the exon 2 product, could give additional substrate specificity, determine the cleavage frame for the alternating NAG–NAM chain and increase the catalytic efficiency of the active centre. Exons 1 and 4, on the other hand, code for protein fragments that are not directly involved in the catalytic function, but may increase the stability of the hen lysozyme molecule and/or carry some as yet unknown biological functions of the enzyme.

The regions of structural homology in T4 and hen lysozyme described by Rossmann⁷ and Matthews⁸ can now be compared with the exon pattern of hen lysozyme defined by Sippel⁶ (see the figure). First, it is clear that residues 1 to 24 or 26 (depending on the comparison) of hen lysozyme, corresponding closely to exon 1 (residues 1–28), have no counterpart in the T4 lysozyme molecule; that is, in T4 lysozyme exon 1 is absent. Second, the principal region of overlap between the two lysozymes corresponds closely to the combination of exons 2 and 3 (residues 28–108). In Rossmann's analysis this region corresponds to residues 25–101 of hen lysozyme and residues 1–73 of T4 lysozyme, while Matthews' analysis⁸ slightly extends and alters the overlap to include residues 27–109 of hen lysozyme and residues 1–80 of T4 lysozyme. Rossmann's analysis⁷ also lines up residues 104–109 of hen lysozyme with residues 100–105 of T4 lysozyme, but this requires a 25-residue insert in the T4 sequence to incorporate it in the major overlap. Accepting some ambiguity with regard to residues 102–109 of hen lysozyme, the correspondence to the combination of exons 2 and 3 is close. If Rossmann's analysis⁷ is restricted to these two exons, the overlap reduces the r.m.s. α -carbon difference from 4.1 Å to 3.2 Å. Finally, correspondence to exon 4 (residues 108–129) of hen lysozyme seems less clear-cut than for exons 1–3. Matthews⁸ does not match it with any part of the T4 lysozyme. Rossmann and Argos⁷ overlap residues 113–126 of hen lysozyme with residues 150–163 of T4 lysozyme but with a rather high r.m.s. α -carbon difference of 6.2 Å. In these circumstances it is difficult to tell whether there is or is not a region of

the T4 lysozyme molecule that corresponds to exon 4 of hen lysozyme.

In conclusion, with the exception of exon 4, the agreement between the exons of hen lysozyme and the structural homology with T4 lysozyme is very close. Exon 1 is absent from the T4 molecule and the principal region of overlap corresponds to exons 2 and 3. In view of Sippel's conclusion⁶ that exon 2 could represent a primitive glycosidase minigene on which substrate specificity was conferred by the product of exon 3, it is tempting to see exons 2 and 3 as the functional nucleus of the lysozyme molecule. The presence of these two regions in T4 lysozyme could account for the similarity in substrate binding and catalytic mechanism suggested by Matthews¹⁰. In addition, the absence of exon 1, and the possible absence of exon 4, are not unreasonable if, as Sippel suggests, they play a role in the stability, rather than catalytic function, of the hen lysozyme molecule. Their stabilizing influence may be taken over in the T4 lysozyme molecule

by the large C-terminal lobe or domain, which has no equivalent in the hen molecule. Indeed, Matthews¹⁰ suggests that this C-terminal region of T4 lysozyme has a function in binding the cross-linked peptide of *E. coli* cell walls, which is not necessary for hen lysozyme.

Although this analysis must be regarded as provisional until other proteins can be examined from a similar point of view, this comparison of T4 and hen lysozyme is fully consistent with the view that recombination within introns can rearrange functional exonic regions into new patterns in new protein products. What is most encouraging is the promise that protein structural analysis can help us identify and predict exons in genes and trace evolutionary events at the level of DNA, and conversely that sequence analysis of gene DNA can help us to understand the structure, function and evolution of the final protein product. □

We are grateful to Professor B. W. Matthews for letting us see his manuscript before publication.

Waves, turbulence, solar convection and ozone

from Michael McIntyre

THE Earth's atmosphere and oceans are increasingly being used as an outdoor laboratory, not only providing information about the behaviour of our environment but also insight into certain difficult, unsolved problems in fluid dynamics. These involve wave-like and turbulent fluid motions at huge Reynolds numbers which are inaccessible to both indoor laboratory experiment and simulation on the world's largest computers.

Empirical information is specially important for the ill understood problem of turbulent motion, and one paper given at a recent conference* was of unusual interest in this respect. J.J. Finnigan (CSIRO, Canberra) and F. Einaudi (Georgia Institute of Technology, Atlanta) have produced what may be the first really detailed information about how an internal gravity wave interacts with pre-existing turbulence. In the case they analysed the wave was nearly monochromatic, and by good fortune had propagated past the 300 m tower and ground-level microbarograph array at the Boulder Atmospheric Observatory, Colorado at a time when a large volume of data was being taken. Turbulent and wave-induced fluctuations could be separated cleanly, revealing how

the turbulent Reynolds stresses varied with the wave motion. To a first approximation the turbulence behaved quasi-elastically, indicating that rapid-distortion turbulence theory should apply. The observed behaviour is quite different from what would have been expected from 'eddy viscosity' or 'mixing length' ideas. It is encouraging to see further direct evidence for the use of rapid-distortion theory, which is complete and rational within its limits of applicability. Other recent lines of evidence come from the work of J.C.R. Hunt and his collaborators at Cambridge on turbulent flow over hills and around buildings, for which rapid-distortion theory is proving to be more useful than previously supposed, and from an estimate by D.O. Gough (Institute of Astronomy, University of Cambridge) using recent observations of the solar 'five-minute' oscillations, showing clearly that they must interact with the Sun's turbulent convection zone in a similar way.

As suggested some years ago by A.A. Townsend, the same ideas seem likely to be important also for understanding the generation of ocean waves by wind. Other aspects of that celebrated and still unsolved problem were illuminated by E. Mollo-Christensen (MIT) and A. Ramamonjia-

1. Gilbert, W. *Nature, News and Views* 271, 501 (1978).
2. Blake, C. C. F. *Nature, News and Views* 273, 267 (1978).
3. Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* 250, 154 (1974).
4. Levine, M., Muirhead, H., Stammers, D. K. & Stuart, D. I. *Nature* 271, 626 (1978).
5. Richardson, J. S. *Biochem. biophys. Res. Commun.* 90, 285 (1979).
6. Jung, A., Sippel, A. E., Grez, M. & Gschütz, G. *Proc. natn. Acad. Sci. U.S.A.* 77, 5759 (1980).
7. Rossmann, M. G. & Argos, P. *J. molec. Biol.* 105, 75–96 (1976).
8. Remington, S. J. & Matthews, B. W. *Proc. natn. Acad. Sci. U.S.A.* 75, 2180 (1978).
9. Matthews, B. W. & Remington, S. J. *Proc. natn. Acad. Sci. U.S.A.* 71, 4178 (1974).
10. Matthews, B. W., Grütter, M. G., Anderson, W. F. & Remington, S. J. *Nature* 290, 334 (1981).
11. Blake, C. C. F. *et al. Proc. R. Soc. B* 167, 378 (1967).

*The American Meteorological Society's third conference on 'Atmospheric and Oceanic Waves and Stability' was held at San Diego from 19 to 23 January, concurrently with AMS meetings on the 'Upper Atmosphere', 'Climate Variations' and the 'Economic and Social Value of Weather and Climate Information'.

Michael McIntyre is in the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge.

risoa (University of Aix-Marseille), who presented new laboratory evidence of strongly nonlinear coupling between the air flow and modulational instabilities of the waves. More light was thrown on the instabilities themselves by a remarkable series of photographs and measurements by Y.-M. Su and A.W. Green (NORDA, NSTL Station, Mississippi) showing a succession of instabilities in a train of moderately steep, monochromatic waves. The waves were generated by a wavemaker about 16m long in an outdoor basin 100 × 340m. A small-scale, essentially three-dimensional instability, very regular in form but as yet unexplained theoretically, was prominent. Its nonlinear evolution appeared to culminate in scattering of some of the wave energy in a direction about 30° away from that of the primary wavetrain, a phenomenon hardly likely to have been identified in a narrow laboratory tank.

Most of our knowledge about the large-scale, low-frequency oscillations known as Rossby waves comes from observations of the atmosphere and ocean, where in practice the waves often have large amplitudes and occupy the no-man's land between 'waves' and 'turbulence'. Experimental possibilities are severely restricted not only by small dimensions but also by the constancy of gravity in the laboratory. Rossby waves can be expected to occur also in the stably stratified interiors of rotating stars. Y. Q. Kang and L. Magaard (University of Hawaii) produced what appeared to be the clearest example yet observed of mid-latitude, oceanic Rossby waves, in a 'baroclinic' mode unobtainable in the laboratory. The wavelength was about 300 km and the waves had a phase velocity towards the north-west and group velocity towards the south-west with magnitudes of the order of a centimetre per second. A point of special interest was that the waves had extremely large amplitudes, as measured by the ratio of a typical particle velocity to a typical phase velocity — well in excess of the kind of amplitudes at which most waves in fluids become very strongly dissipative.

Rossby waves in the atmosphere, too, were a topic of renewed interest. A number of papers were on theoretical models related to the recent work of J.G. Charney (MIT) and his collaborators, in which global-scale, topographically generated Rossby waves are allowed to interact with an azimuthal mean flow driven by a constant forcing. This simple interaction seems capable of explaining the observed tendency of the atmospheric circulation to flip between quasi-stable states, with large and small amplitudes in those wave components which are phase locked to the topography. The large-amplitude states are associated with 'blocking' patterns and persistent spells of anomalous weather. In winter the same Rossby waves can penetrate the stratosphere, where under the right conditions they trigger 'sudden

warmings' — sporadic but spectacular upheavals in the state of the polar stratosphere in which dynamically induced subsidence of air over the polar cap causes large, adiabatic temperature rises. Whether or not a sudden warming occurs is sensitively dependent on conditions in the stratosphere itself, as well as in the troposphere. At the concurrent meeting on the 'Upper Atmosphere' there were signs that our understanding of the conditions favouring sudden warmings may be about to improve significantly. Recent progress in the dynamical theory, together with the uniform global coverage offered by infrared satellite observations, have led to better ways of analysing the observations

and more illuminating comparisons with computer simulations. Some real clues are emerging as to how and why the behaviour of the stratosphere differs from that of the computer simulations, and we can await the results of current modelling efforts with considerable interest. The sudden warming is a large-scale dynamical experiment which nature kindly performs from time to time. It is one of the crucial tests of our ability to simulate the large-scale dynamics of the stratosphere. The dynamics, in turn, are one of the necessary ingredients in attempts to understand the stratospheric circulation in general, and to predict the effects of pollution on the ozone layer in particular. □

Climatic variation in San Diego

from M.J. Salinger and C.B. Sear

DURING anomalously warm and dry weather in San Diego in January, climatologists met to discuss climatic variations. Much of the meeting* centred on the implications for global climate of the likely increase in atmospheric CO₂ concentration during the next few decades, but other topics, including recent climatic change and palaeoclimatic studies, were also discussed and will be highlighted here.

A new Southern Hemisphere data set, covering 1972 to 1979, has recently become available from the Australian Bureau of Meteorology and several aspects of the circulation down under were discussed, using these data. M.L. Blackmon and T.J. Phillips (National Center for Atmospheric Research (NCAR), Boulder, Colorado) examined the 1,000 mbar height data and showed that, between 1972 and 1979, there was a band of high variance between 40° and 60°S from south of Africa to New Zealand. Together with the results of K.E. Trenberth¹ which show a similar band of maximum variance in mid-tropospheric data at 500 mbar, these investigations suggest that the circulation of the Southern Hemisphere consists essentially of one major storm track. Corroborative evidence also came from J.C. Rogers (Ohio State University) who used Eigenvector analyses to show that negative correlations exist between high and middle latitudes in the same longitudinal band during the Southern Hemisphere summer. These results suggest that, when the westerlies in one sector of the hemisphere are strong, they are also strong in other sectors.

Chiu and Newell (MIT) described similar analyses for the Northern Hemisphere using upper air data for the period 1959–79 and concluded that the non-seasonal low-level changes during this period were dominated by the 'Greenland see-saw'

which has been described previously, notably by H. van Loon and J. Williams². The see-saw operates when surface variables, pressure and temperature, in the region of southern Greenland and northern Scandinavia are negatively correlated. At higher levels (200 mbar) Chiu and Newell found that a persistent pattern, centred over Greenland, dominated. They suggested that this could be related to the Southern Oscillation which is a quasi-periodic four to six-year climate oscillation centred over the Pacific and linked to such important phenomena as El Niño.

First described as early as 1936 by Walker and Bliss, the Southern Oscillation is receiving a lot of attention because it is one of the best defined quasi-regular climatic events. Although primarily a tropical Pacific phenomenon, it has often been proposed as an important controlling mechanism for variations over much of the globe. Indeed, the studies described in San Diego confirmed old and elucidated new 'teleconnections' between the tropical Pacific and higher latitudes in both hemispheres. K.E. Trenberth and D.A. Paolino (University of Illinois) reported the variability of the Northern Hemisphere surface circulation showing connections with the Southern Oscillation. They used data from the period 1925–77 and Empirical Orthogonal Function (EOF) analyses to show that the dominant pattern in all seasons was one associated with anomalies centred on the Aleutian and Icelandic low-pressure areas. The time series of this pattern have significant periodicities at about six years which, they speculate, are associated with the Southern Oscillation.

Of the researchers who discussed synoptic analyses, only D.J. Shea (NCAR)

M.J. Salinger and C.B. Sear are at the Climatic Research Unit, University of East Anglia.

*The first AMS conference on 'Climate Variations'. See footnote on previous page.

presented results from global data using EOF analyses. His sensitivity studies suggested that the first three EOFs of temperature were stable but that none of the EOFs of precipitation is so, implying a very high degree of randomness in spatial patterns of precipitation. This is an important consideration, particularly for those who are searching for a CO₂-induced climatic change and are finding analogues from within the recent period.

Seemingly far removed from such studies were those using palaeoclimatic data to investigate long-term climatic change. Studies of deep-sea cores have provided considerable insight into climatic variations on such time scales. W.C. Prell (Brown University) concluded from examination of foraminifera assemblages from the western Arabian Sea that a distinctive assemblage occurs when the summer monsoon causes intense upwelling. The sediments provide a record of monsoonal upwelling variations over the past 150 thousand years which suggests that the monsoon was more intense from about 8,000 to 10,000 years ago — a time of increased precipitation in Africa and Arabia³. The more intense monsoon coincides with the dates of maximum Holocene warming reported in Australia and New Zealand. There is now mounting evidence that the most pluvial and warmest periods of the past 15,000 years did not occur simultaneously throughout the globe and this should be borne in mind by those attempting to produce analogues for a warmer, high-CO₂ world using data from the Holocene.

A model of oceanic moisture feedback to ice sheets in the Northern Hemisphere was advanced by W.L. Ruddiman and A. McIntyre (Lamont-Doherty Geological Observatory, Columbia University; see ref. 4). They have found evidence from a core in the central North Atlantic that, for the past 250 thousand years, North Atlantic moisture feedback has amplified Milankovitch forcing, defined in summer insolation, from the 23,000-year precessional cycle, and have postulated a feedback mechanism in which orbitally controlled changes in summer insolation over the North American ice regulate meltwater/iceberg outflow to the North Atlantic. The feedback is such that when summer insolation is high, ice rafting into the ocean causes the formation of a low-salinity layer at the surface. This in turn restricts moisture transport to the ice sheet. The opposite happens when insolation is low. Accelerated shrinkage or growth

occur accordingly and positive feedback results.

Among the many modelling studies discussed were some concerned with palaeoclimatic circulation. A low-order general circulation model (GCM) of the Earth's atmosphere, developed by B.L. Otto-Blieser (University of Wisconsin-Madison), was used by J.E. Kutzbach (at the same institution) and T. Webb III (Brown University) to simulate the climate 18,000 years ago. The rationale behind the use of a low-resolution model was that palaeoclimatic data have similar resolution. If the model is successful, then both modern data and data from 18,000 BP can be used to simulate climate in the intervening epochs. Their simulation of the climate of 18,000 BP had an enhanced westerly gradient across the US with the strongest flow just south of the ice sheet, in agreement with earlier GCM results⁵.

S.L. Thompson, E.J. Barron and S.H. Schneider (NCAR) reported energy balance modelling studies, based on the work of Barron *et al.*⁶, which investigated the causes of the ice-free mid-Cretaceous climate. Their results suggest that changes in zonal palaeogeography (continental drift) are an important forcing mechanism for climate change but that meridional temperature distribution cannot be simulated successfully without additional feedback assumptions. Their model included albedo and cloud feedbacks and meridional heat transport changes but could still not model the Cretaceous temperature distribution. Thompson and his colleagues speculated that to attain Cretaceous levels of warming, the Earth could require additional external forcing from the Sun. Although indirect, this is one of the first quantitative investigations implicating solar constant changes as causes of climatic change on the time scale of hundreds of millions of years.

Few new results were presented on the recurrent theme of carbon dioxide, but modellers of a CO₂-warmer Earth did not get away from San Diego without warnings, notably from S.H. Schneider and S.L. Thompson that a CO₂-warmer Earth has so far been modelled using an instantaneous increase in CO₂ concentration, whereas, of course, in the real world the concentration increase occurs very slowly⁷. Thompson and Schneider suggest that the thermal inertia of the upper ocean, together with vertical mixing of the deep ocean waters, could delay the response of the globally averaged surface temperature to increasing CO₂ by a decade or so relative to equilibrium calculations. They suggest further that the evolution of surface climate anomalies, as CO₂ increases, may be significantly different from that suggested by equilibrium modelling results. They stress that the nature of this transient response is a major uncertainty in present CO₂-climate change studies and that further research into this topic is urgently required. □

The return of Halley

from David W. Hughes

HALLEY'S comet was last seen at the end of May 1911 when it was at about Jupiter's distance from the Sun and heading out into the Solar System beyond the orbit of Neptune. Aphelion was reached in 1948. The comet has been on its way back ever since.

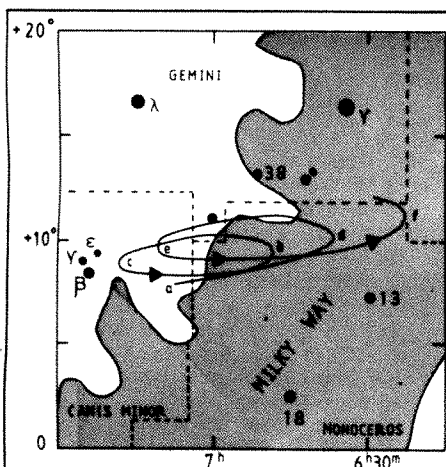
The search for Comet Halley began, rather optimistically, in November 1977 when observers at two of the United States' largest telescopes unsuccessfully attempted to recover it more than eight years before the predicted perihelion passage date of 9 February 1986. The comet's estimated magnitude was fainter than 26 at that time.

The early recovery of the comet is of considerable importance as this would enable a valuable check to be made on the predicted orbital parameters. Comets are much affected by the gravitational perturbations induced by the major planets. In the case of Halley this causes a slow variation in the period, between extreme values of about 76.1 yr and 79.4 yr, and a continuous increase in the longitude of the nodes. Spacecraft missions, like the planned Giotto Comet Halley flyby, require an accurate knowledge of the comet orbit to enable the spacecraft to fly as close to the comet as possible. At the last apparition the comet was recovered by Max Wolf using a 1 hour photographic exposure with the 72 cm Heidelberg reflecting telescope. The comet appeared as a small fuzzy image of magnitude 16. It was 3.4 AU from the Sun, in the outer regions of the Asteroid belt and 7.5 months from perihelion passage. At the present time the comet is 14 AU from the Sun and has an estimated magnitude of 24.7. (A change of 1 in magnitude is equivalent to a brightness ratio of about 2.5, so a comet of magnitude 16 is 3,000 times brighter than a comet of magnitude 24.7.)

The estimated path of Comet Halley is shown in the accompanying figure. Data for this figure have been extracted from *The Comet Halley Handbook: an Observer's Guide* which has been produced for the International Halley Watch by Donald K. Yeomans of the Jet Propulsion Laboratory, Pasadena, California (JPL publication 400-91 1/81). The comet is near the boundary of the constellations Canis Minor and Monoceros. The fact that the Milky Way acts as a backdrop obviously makes the search more difficult. The loopy path of the comet against the sky background is caused mainly by the movement of the observer. The comet is

David W. Hughes is in the Department of Physics, University of Sheffield.

1. Trenberth, K.E. *Mon. Weath. Rev.* (in the press).
2. van Loon, H. & Williams, J. *Mon. Weath. Rev.* **104**, 366 (1976).
3. Street & Grove *Nature* **261**, 385 (1976).
4. Ruddiman, W.L. & McIntyre, A. *Science* **204**, 173 (1979).
5. Williams, J. in *Climatic Change* (ed. Gribbin, J.) (Cambridge University Press, 1978).
6. Barron, E.J. *et al. Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30**, 17 (1980).
7. Schneider, S.H. & Thompson, S.L. *J. geophys. Res.* (in the press).



The predicted path of Halley's comet for the next two years. The comet is mainly in the constellations Canis Minor and Monoceros. a, b, c, d, e and f represent its positions during January 1981, April 1981, October 1981, April 1982, October 1982 and April 1983 respectively. In January 1981 the comet had an estimated nuclear magnitude of +24.8 and this brightens steadily to about +22.5 in April 1983.

seen at maximum right ascension in October and at minimum right ascension in April. In this time interval the Earth has moved about 300,000,000 km, from one side of its orbit to the other. As the comet approaches the Sun the size of the loop increases. The most favourable time of the year for recovery will be when it has its maximum rate of angular movement against the sky background and can thus be more easily distinguished from a star. This occurs during the horizontal portions of the curve in the figure, at times roughly corresponding to winter and summer.

The fame of Halley's comet is not due to its brightness. Indeed, many of the accounts we hear from parents and grandparents of the appearance of the comet in 1910 do not refer to Halley at all but to the Daylight Comet, 1910 I, also known as the Great January Comet. Unfortunately, the forthcoming apparition will be unfavourable. Many of the waiting public will be disappointed as the comet will not be obvious to the naked eye. When Halley is at its most active it will be on the opposite side of the Sun to Earth and its maximum magnitude will be only about 2.9, about 2.5 times fainter than Polaris, the pole star.

Yeomans reviews the observing prospects and provides a series of charts showing the elevation, azimuth and magnitude of the comet in the morning and evening sky for latitudes 40°N, 30°N, 20°N, 20°S and 30°S. Two things are only too clear. Southern Hemisphere observers are better placed than Northern Hemisphere ones and people living at latitudes above 50°N are going to have considerable problems. It will be worth the effort though. One thing is certain, the similarity between our 'three score years and ten' and Halley's orbital period makes this a once in a lifetime experience for the large majority of us. □

Hybrid myosins

from Dixie W. Frederiksen

A recent paper by Sellers, Chantler, and Szent-Gyorgyi¹ of Brandeis University has provided important new insight into the control of contraction in molluscan muscle and into the properties of myosins from a wide variety of species and cell types. It is the interaction of actin with myosin that brings about activation of the myosin Mg-ATPase and the production of mechanical work for contraction or cell motility². The intracellular concentration of free calcium controls this interaction by mechanisms which differ in different systems.

In vertebrate striated muscle the primary effect of calcium is on troponin, a protein bound to actin and to tropomyosin on the thin filament of the sarcomere³. Conformational changes induced by calcium lead to a removal of inhibition of the actin-myosin interaction. In the striated muscles of molluscs and some other invertebrates, calcium acts directly on myosin in the thick filament to induce conformational changes that allow myosin to bind to actin⁴. In smooth muscle the situation is less well understood but the primary action of calcium appears to be an activation of a myosin light chain kinase. This produces a phosphomyosin which, unlike the unphosphorylated protein, can be activated by actin⁵. Additional effects of calcium have been reported for vas deferens and vascular smooth muscle⁶ and it has also been shown that smooth muscle contraction may be calcium regulated through an actin-linked system⁷.

In all muscle and in most non-muscle cells, the myosin molecules⁸ are hexameric species that consist of two 'heavy' (200,000 molecular weight) and four 'light' (14,000–31,000 molecular weight) polypeptide chains. There are two classes of light chains; the two 'non-essential', or 'regulatory', light chains are relatively easily removed from the major portion of myosin without loss of biological activity while the two 'essential' or 'alkali' light chains are removed from myosin only after rather harsh treatment with alkali or denaturant and their loss is accompanied by loss of ATPase activity. The intact myosin molecule is highly asymmetric with helical portions of the heavy chains contributing most to the length of the molecule. The light chains and globular portions of the heavy chains comprise a pair of structures at one end of the molecule that each contain an actin-binding and an ATPase site.

The detailed structures and functions of the myosin subunits vary from one cell type to another and, in their recent paper, Szent-Gyorgyi and his colleagues illustrate some important similarities and distinc-

tions among the various regulatory light chains from different myosins. Using some of the techniques they developed earlier⁹, they prepared approximately twenty different hybrid myosins containing the heavy and alkali light chains of scallop adductor myosin and two regulatory light chains from other myosins. They then tested regulatory light chains from vertebrate and invertebrate striated muscle, gizzard smooth muscle and platelets. Since all of these foreign subunits bind easily to light chain-depleted scallop myosin, the heavy chain-binding sites on all the regulatory light chains must be similar, as had been previously suggested by Weeds and his collaborators¹⁰.

The properties of the hybrid myosins are consistent with the idea that smooth muscle, as well as molluscan and some other invertebrate striated muscle myosins, are calcium-sensitive regulatory switches for the contractile process, but the vertebrate striated muscle has no myosin switch. Not surprisingly, the regulatory light chains from scallop and other molluscs restore full calcium ion sensitivity to the scallop myosin. In addition, light chains from crab, cricket, gizzard and platelets also form hybrids that bind calcium and require the divalent cation for actin activation of the ATPase. (The concentration of free calcium necessary for activation of these hybrids is higher than that required for the molluscan hybrids.) Vertebrate and lobster striated muscle myosin light chains produce calcium-insensitive hybrids without high-affinity calcium binding sites. Phosphorylation of the vertebrate striated and smooth muscle light chains (see *Myosin phosphorylation* *Nature* 287, 191; 1980) did not affect the enzymatic properties of the hybrid myosins.

The Szent-Gyorgyi group has also resolved some of the puzzling questions about the mechanism of calcium control in molluscan muscle. Because the first regulatory light chain is much more easily removed from scallop myosin than the second, it seemed possible that the two light chain binding sites on scallop myosin are not identical. In an elegant series of experiments using labelled regulatory light chains, they showed that the light chains are removed in an order independent of the order of combination. The subunits bind to the main portion of myosin in a cooperative manner. The cooperativity is negative,

1. Sellers, Chantler & Szent-Gyorgyi *J. molec. Biol.* 143, (1980).
2. Adelstein & Eisenberg *A. Rev. Biochem.* 49, 921 (1980).
3. Weber & Murray *Physiol. Rev.* 63, 612 (1973).
4. Szent-Gyorgyi *Biophys. J.* 15, 707 (1975).
5. Dabrowska *et al. Biochem.* 17, 253 (1978); Hathaway and Adelstein *Proc. natn. Acad. Sci. U.S.A.* 76, 1653 (1979).
6. Chacko *et al. Proc. natn. Acad. Sci. U.S.A.* 74, 129 (1977); Rees & Frederiksen *J. biol. Chem.* 256 (1981).
7. Mikawa *et al. J. Biochem. (Tokyo)* 82, 1789 (1977).
8. Mannherz & Goody *A. Rev. Biochem.* 45, 427 (1976).
9. Kendrick-Jones *et al. J. molec. Biol.* 104, 747 (1976); Chantler & Szent-Gyorgyi *J. molec. Biol.* 138, 473 (1980).
10. Weeds *et al. in Calcium Binding Proteins and Calcium Function* (eds Wasserman *et al.*) 222 (Elsevier North-Holland, Amsterdam, 1977).

Dixie W. Frederiksen is in the Department of Biochemistry, Vanderbilt University, Nashville, Tennessee.

however, since the affinity of myosin for light chain drops dramatically after the first one is bound. The affinity of light chain-depleted scallop myosin for foreign light chains is as great or greater than the

affinity for scallop light chains. Although this may be a fortuitous property of scallop myosin, it is intriguing to ask what might be learned using other light chain-depleted myosins in similar experiments. □

Two distinct types of oceanic plateau

from Peter J. Smith

THE many refraction studies carried out over the oceans during the past twenty years have shown the basic structure of the oceanic crust to be remarkably uniform. The three-layer model proposed by Raitt (in *The Sea*, Wiley, 1963) in pre-plate tectonic, and almost pre-ocean floor spreading, days still generally holds good, notwithstanding the discovery of systematic variations in some lithospheric properties with age. At the same time, however, the ocean basins contain some obvious anomalous features, of which active ridges are the most conspicuous and hot spots the most controversial. And then there are the 'ocean plateaus' — broad aseismic platforms standing high above the ocean floor and usually topped with thick deposits of calcareous sediments.

Some of these plateaus (Rockall Bank and Seychelles Plateau, for example) have long been regarded as continental fragments, presumably broken off from larger landmasses during continental redistri-

bution. But others (for example, Broken Ridge and Shatsky Rise) are thought to be truly oceanic, although the processes by which they are formed are unknown. To test this two-fold classification and put it onto a more systematic basis, Carlson *et al.* (*Earth planet. Sci. Letts* **51**, 171; 1980) have now collated the seismic refraction data from 11 of the many oceanic plateaus known, showing that the two types are structurally distinct as might well be inferred from their supposedly different origins.

Of the 11 features studied, three were already known to be underlain by rock of continental type. Thus Rockall contains Lower Eocene granites; granitic rocks are exposed on the Seychelles Plateau; and gneissic rocks have been recovered from the basement of the Falkland Plateau by the Deep Sea Drilling Project. In addition,

Saya de Malha Bank is apparently part of the same physiographical feature as the Seychelles, the Mozambique Plateau is close to Africa and thought to be a piece broken from it, and Lord Howe Rise is likewise thought to be a fragment from nearby Australia, although in these three cases there is apparently no direct evidence of the presence of continental material.

But direct evidence or not, the seismic velocity-depth profiles suggest that these six platforms are of similar construction and thus members of a class. In three cases (Seychelles, Rockall, Lord Howe) upper mantle velocities ($>8.0 \text{ km s}^{-1}$) are observed at about 30 km below sea level, and immediately above the Moho there is a 14–18 km thick lower crustal layer with seismic P wave velocities in the range $6.0\text{--}7.0 \text{ km s}^{-1}$. Unfortunately, the profiles from the three other plateaus (Mozambique, Saya de Malha, Falkland) do not extend down to the Moho, although the same basal crustal layer is present in Mozambique and Saya de Malha. Still more unfortunately, the Falkland profile does not even reach that far, but is consistent with the others in having a wide range of velocities of up to 6.4 km s^{-1} in its upper parts.

This upper-level velocity variability is also evident in eight seismic profiles from the five platforms suspected not to be of continental origin. But there the resemblance ends. Generally (five profiles — Shatsky, Broken Ridge, Fiji) the mantle is reached at depths of only 18–25 km, and immediately above the Moho there is a lower crustal layer 6–15 km thick having seismic velocities in the range

Peter J. Smith is the Editor of 'Open Earth'.



THE FALLS OF NIAGARA IN WINTER

In the first week of last February it fell to my lot to make very hurriedly the transcontinental journey of 3500 miles from San Francisco to New York. Before starting I resolved that the one stoppage which I could allow myself *en route* should be made at Niagara. I was confirmed in my resolve by seeing statements in various American papers to the effect that, owing to the long-continued and exceptionally severe cold of the present winter, the Ice-mountains at the Falls were higher than had ever been previously known.

In the Niagara gorge numerous springs discharge themselves into the chasm at various points in the precipitous rocky sides, and at these points numerous collections of huge and massive icicles appeared as though adherent to the rock, measuring perhaps seventy or eighty feet in length, and eight to ten feet in irregular diameter.

The boulders in front of the American Fall are, as it were, gigantic nuclei, round which the frozen spray accumulates, and produces the

Ice-mountain of which we hear so much, and the remains of which are not unfrequently to be seen even by summer visitors. The average height of this is about half the total height of the Fall, but this winter it has attained to the unprecedented height of within twenty feet of the top of the Fall!

Within the last few years a considerable portion of "Table Rock" has fallen away. In its present condition a stream of water about one foot in thickness falls over it in summer, and, owing to the amount of its overhanging, it is easy to get between this Fall and the rock, and thus to be "behind Niagara." At the time of my visit (February 8th), however, the whole of this portion of the Fall was *completely frost-bound*. Enormous icicles, of the most surpassing beauty, depended from the rock above, while at my feet were masses of the frozen spray from the Horse-Shoe Fall. The intense emerald green of the water of that Fall, seen through and between these magnificent ice-pendants, could be reproduced by no artist, but will never be effaced from my memory. The accompanying woodcut, photographed on to the wood block from a photographic picture taken a few days prior to my visit, will, to those who know the place, give some faint idea of the beauty of the scene, and of the gigantic scale of the icicles.

From *Nature* **23**, 31 March, 511, 1881.



Gigantic icicle under Table Rock. The upper right-hand corner is rock; a portion of the Canadian (or Horse-Shoe Fall) is on the left. The whole of the apparent ground is a mass of frozen spray which has accumulated many feet in thickness on the shingle at the foot of the rock.

7.3–7.6 km s⁻¹. Then above that there is a unit with typical oceanic layer 3 velocities. The only exception to all this (apart from the Manihiki Plateau whose profile is too short to indicate much) is the Ontong–Java Plateau, the two profiles of which demonstrate a curious mixture of both ‘continental’ and oceanic characteristics. On the face of it, this spoils the story rather, although it could quite genuinely represent a composite structure with both continental and oceanic elements.

Leaving that anomaly aside, however, there are clearly just two basic and distinct structural patterns, suggesting two separate modes of origin. The two types of structure also correlate well with their respective counterparts elsewhere. Thus the first (continental) pattern is very similar to the seismic velocity–depth structure of continental shields, whereas the second (oceanic) is not. Then, the high basal

velocities evident in the second pattern have also been observed at the base of normal crusts in parts of the Pacific and the Atlantic, as well as in some ophiolites, which in fact strengthens the view that oceanic plateaus of the second kind are genuinely of marine origin.

Carlson and his colleagues admit that the number of plateaus for which velocity profiles are available is small and therefore regard their analysis as tentative, although the agreement between the seismic data and prior suppositions about two different origins inspires confidence. The mode of formation of the non-continental plateaus remains open, however. The new analysis seems to confirm the conclusion by Hussong and his colleagues (*J. geophys. Res.* **84**, 6003; 1979) that the truly marine plateaus are expanded sections of normal oceanic crust, but the cause of the expansion is a mystery. □

lymphocytes from many long-standing MG patients make anti-AChR in culture, it seems that the thymus is not the primary site of anti-AChR synthesis (J. Newsom-Davis, Royal Free Hospital, London). Thymic lymphocyte preparations (irradiated to destroy B cells) can also enhance the *in vitro* anti-AChR synthesis by MG peripheral blood lymphocytes, possibly due either to antigenic stimulation by AChR on thymic cells, or to the presence of a thymic helper-T cell population specific for anti-AChR (or both) (Newsom-Davis). Other cell types may also be involved, however, as N. Shinomiya (Tokyo Medical and Dental School) showed that normal, but not MG, peripheral blood T cells inhibited anti-AChR synthesis by MG lymphocytes. This suggests the absence of a specific suppressor-T cell in MG subjects.

This latter finding indicates that even normal subjects may have AChR-sensitive lymphocytes and display low levels of anti-AChR. Anti-AChR has been detected in the sera of 18 per cent of MG relatives (R. Pirskanen, Karolinska Institute) although it did not correlate with HLA-B8, the HLA antigen most closely associated with MG, nor with the electromyographic abnormality also found in a proportion of MG relatives. Others have demonstrated very low levels of anti-AChR in normal subjects (T. Mittag, Mount Sinai School of Medicine), in cultures of normal lymphocytes with MG T cells (R. Kelley, UCLA School of Medicine) and in purified Torpedo AChR (Shinomiya, Tokyo Medical and Dental School). The existence of sensitized B lymphocytes would not be surprising if the normal immune system was exposed to AChR, and it may be relevant that AChR-like molecules have been detected on normal peripheral lymphocytes (G. Morrel, Wayne State University; D. Richman, University of Chicago) and mouse thymic lymphocytes (S. Fuchs, Weizmann Institute) although α -bungarotoxin binding has not been demonstrated.

Work on EAMG may help to elucidate the involvement of genetic factors. When animals are immunized against AChR, a proportion of the anti-AChR antibodies cross-react with the animals’ endplate AChR, resulting in weakness and neuromuscular block. Not all species or strains respond equally and P. Berman (Salk Institute) showed that susceptibility to clinical EAMG in inbred strains of mice was not related to the anti-AChR titre, its ability to increase AChR degradation, or a difference in the safety factor for transmission at endplates from different genetic strains. The greatest susceptibility was found in those with a particular major histocompatibility complex (H-2^b) and immunoglobulin (IgT^b) genotype. The involvement of the IgC_H region suggests that overt myasthenia in mice depends on particular polymorphic genes coding for the constant region of immunoglobulin

Idiotypic restriction in myasthenia gravis antibodies

from Angela Vincent

It is now reasonably certain that anti-acetylcholine receptor antibodies (anti-AChR) play a pathogenic role in the reduction in endplate acetylcholine receptors (AChR) that underlies the defective neuromuscular transmission characteristic of myasthenia gravis (MG). Experimental autoimmune MG (EAMG) in animals immunized against purified AChR, human anti-AChR in over 80% of MG patients, and passive transfer of MG to mice using MG immunoglobulins had all been reported by 1975 and progress made since then was the subject of a recent conference*.

It has been shown that, although in any individual the initial anti-AChR titre gives no indication of the severity of the disease, subsequent anti-AChR levels tend to correlate with changes in the clinical state. Various mechanisms thought to be involved in loss of AChR function have been demonstrated, including complement-dependent damage to the postsynaptic membrane (A. Engel, Mayo Clinic), increased AChR degradation caused by cross-linking of AChR by divalent antibody (D. Drachman, Johns Hopkins University), and direct ‘block’ of AChR function by binding of antibody to the ACh/ α -bungarotoxin binding site (D. Drachman; B. Fulpius, University of Geneva). Antibodies can also modify rather than block AChR function, as E. Albuquerque (University of Maryland)

showed that a factor in MG serum can slightly increase the open lifetime of the individual AChR-associated ion channels. This effect could only be seen, however, when the serum had been applied in the presence of histrionicotoxin, a toxin that reversibly blocks the channel in its open state.

Although the pathogenesis of the disease appears to be established, the fundamental aetiology and cellular immune mechanisms are not. Is there a precipitating event such as virus infection, and how are genetic factors involved? Is there a specific lymphocyte abnormality or a failure of normal control mechanisms? Any answer to these questions must include an explanation of the role of the thymus. Many patients have ‘thymitis’ with germinal centres, and some have a thymic tumour. MG symptoms, particularly in the former group, often respond well to thymectomy although improvement may take months or years. One possibility is that the thymus is the site of both the antigen and the initiating event. It has been recognized for several years that cultured thymic epithelial cells can display muscle-like antigens including AChRs (K. Wekerle, Max Planck Institute). Wekerle proposed that helper-T cells were sensitized to AChR in the thymus and then emigrated to the periphery where anti-AChR synthesis took place. This may well be the case because, although thymic

Angela Vincent is in the Department of Neurological Science, The Royal Free Hospital, London.

*The sixth conference on myasthenia gravis was held in New York from December 2nd to 4th, 1980, organized by Dr D. Grob (Maimonides Medical Center) and sponsored by the New York Academy of Sciences and the Myasthenia Gravis Foundation.

heavy chains. A similar association (with the human Gm allotypes) has been reported in MG (Nakao *et al. Lancet*, 677; 1980). These findings might explain the generally poor correlation between the total anti-AChR titre and disease severity. Clearly, specific subpopulations of antibodies may also be important in MG and will be defined when human monoclonal anti-AChR is available.

In EAMG several workers have used monoclonal antibodies produced by hybridomas to analyse the anti-AChR. A single monoclonal antibody reactive with mammalian AChR can passively transfer EAMG to mice (J. Lindstrom, Salk Institute) or rats (V. Lennon, Mayo Clinic). Surprisingly, antibodies raised in inbred rats against each of three distinct monoclonal antibodies inhibited the activity of the others suggesting that they shared a common idio type (the antigen-specific, hypervariable region) (V. Lennon). However, each anti-idiotypic antibody inhibited only a proportion of the anti-AChR in polyclonal EAMG sera and

when anti-idiotypic-producing animals were immunized directly against AChR they still developed electrophysiological evidence of EAMG.

J. Lindstrom has used monoclonal antibodies raised against intact AChR and its polypeptide subunits to map determinants on receptor extracted from many different species. The main immunogenic region appears to be on the α -subunit of Torpedo and eel AChR and similar determinants are found on AChR from mammals including humans. Three monoclonal antibodies reacting with α , β and γ subunits can bind to human AChR and inhibit the binding of a proportion of human anti-AChR (S. Tzartos, Salk Institute). These results are important because they suggest that the determinants to which human anti-AChR bind are more restricted than was previously inferred from the heterogeneous characteristics of MG anti-AChR (as defined by variability in κ and λ light chains, binding to the α -bungarotoxin binding site, cross-reactivity with other AChR preparations).

Moreover, using a polyclonal anti-idiotypic serum raised in rabbits against human anti-AChR, A. K. Lefvert (Karolinska Institute) has found some sharing of anti-AChR idiotypes between different MG patients.

The treatment of MG has improved since the autoimmune basis for the disease has been widely recognized. Immunosuppressants are used much more freely with good results, and plasma exchange can produce marked temporary improvement. Although several interesting experimental approaches to treatment, including immunization against modified AChR and anti-idiotypic induction in rabbits with EAMG (S. Fuchs), have met with some success, it will be some time before any can be tried in the human disease. Nevertheless, if there is less heterogeneity and greater idio type restriction in MG anti-AChR than previously thought, and if certain idiotypes can be shown to play a crucial role in the disease process, specific immunotherapy may become a realistic goal.

Disposing of dioxins by oxidation

from Alastair Hay

DISPOSAL of waste contaminated by polychlorinated dibenzodioxins (PCDDs) has long been a problem for sections of the chemical industry. Toxic and often remarkably stable, the PCDDs are produced as unwanted by-products during the manufacture of chlorinated phenols.

The best known and most extensively studied PCDD is 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). It is produced during the manufacture of 2,4,5-trichlorophenol — the starting point for the widely used herbicide 2,4,5-trichlorophenoxyacetic acid (2,4,5-T).

Although pure TCDD is very stable (a temperature in excess of 800°C is required for its thermal degradation), if dissolved in organic solvents such as methanol or benzene it is exceptionally unstable in the presence of ultraviolet light. The solvents act as hydrogen donors, replacing aromatic ring chlorines with hydrogen as TCDD is photolytically degraded. The reaction sequence proceeds in stages with only one chlorine atom being replaced at any one time. Eventually TCDD is degraded to the unchlorinated parent dibenzodioxin (D.G. Crosby *et al. Science* **173**, 748; 1971). Many orders of magnitude less toxic than TCDD, dibenzodioxin too is decomposed by light (J.R. Plimmer *et al. Adv. Chem.* **120**, 44; 1973).

The higher chlorinated dibenzodioxins, hexa-, hepta- and octachlorodibenzo-*p*-dioxin, produced in the manufacture of the wood preservative pentachlorophenol, are also subject to photolytic degradation (C. Rappe *et al. Ann. N.Y. Acad. Sci.* **320**, 1;

1979).

Under field conditions TCDD can be subjected to photolytic degradation by sunlight when a suitable solvent/hydrogen donor is present. When a preparation of the herbicide 'Agent Orange' — a formulation of 2,4,5-T containing 15 p.p.m. TCDD — was exposed to sunlight, the dioxin was rapidly degraded with less than half remaining after 6 hours (A.S. Wong and D.E. Crosby in *Dioxin: Toxicological and Chemical Aspects* (eds Cattabeni, F. *et al.*, Spectrum, New York; 1978). When TCDD was dissolved in an oil/emulsifier mixture and exposed to sunlight, complete degradation occurred within 2 days (A.S. Wong and C.G. Crosby *op. cit.*).

In industry, photodegradation of PCDDs is not very practical and most manufacturers dispose of the chemicals by high-temperature incineration. A method of degradation using chemical processes might provide a more attractive option and in this issue of *Nature* (p.323), D.C. Ayres reports that PCDDs can be degraded efficiently by an oxidative procedure involving ruthenium tetroxide.

Ruthenium tetroxide is a powerful oxidizing agent which reacts with a wide range of organic substances, and according to Ayres, it can be used in solution in water or in organic solvents with no nucleophilic character, such as chloroform, nitro-

methane or carbon tetrachloride. Ayres reports that 2,7-dichlorodibenzo-*p*-dioxin (DCDD) was destroyed by ruthenium tetroxide in carbon tetrachloride solution when the cocktail was allowed to stand overnight at ambient temperature. The rate of oxidation increases with increase in temperature: at 30°C DCDD has a half life of 215 min whilst at 50°C, the half life falls to 38 min. The oxidation of TCDD is a slower process: at 20°C this chemical has a half life of 560 min in the ruthenium tetroxide mixture. At 70°C, however, the half life plummets to a mere 15 min.

Ayres believes that his results demonstrate that ruthenium tetroxide can be safely used for the detoxification of glassware and the periodical purging of industrial reactors to counteract the accumulation of PCDD residues. The process need not be expensive, Ayres claims, for ruthenium tetroxide need only be used in catalytic amounts after which it can be regenerated using a consumable secondary oxidant, such as hypochlorite.

Although Ayres does not know the nature of the fragments formed during the oxidation of the PCDDs, he reports that chlorophenols subjected to ruthenium oxidation undergo extensive decomposition to yield chloride ions and no significant levels of organic products. It remains to be seen whether this is the fate of the PCDDs, but it is clear that oxidation must now challenge incineration as the method of choice for disposing of unwanted polychlorinated dibenzo-*p*-dioxins in industry.

Alastair Hay is in the Department of Chemical Pathology at Leeds University.

ARTICLES

Orbital signature of interglacials

G. Kukla*, A. Berger†, R. Lotti* & J. Brown*

* Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA

† Institute of Astronomy and Geophysics, Université Catholique de Louvain, Louvain-La-Neuve, Belgium

A specific orbital configuration—high obliquity combined with the June perihelion—marked the beginning of the past three interglacials. This suggests that the primary cause of the glacial cycle may be astronomical. An astronomical climate index (ACLIN) is introduced which combines the three orbital variables in the time-lag bivariate model designed to predict the major climate changes in the late and middle Pleistocene, and in the near future. ACLIN closely correlates with the major climatic events revealed by independently dated proxy climate indicators of the past 130,000 yr. It successfully differentiates the interglacials, and displays a 100,000-yr periodicity. It predicts an early end of the present interglacial and the start of a new one in 114,000 yr.

SEVERAL recent observations have shown that the principal periods in the radiometrically dated palaeoclimatic records match in length the periodicities of the Earth's orbit¹⁻⁵. These two systems are probably linked by a cause and effect relationship. Although there have been several attempts to explain the palaeoclimatic record in terms of insolation input, none of the selected palaeoinsolation curves correlates with the different amplitude of interglacials and interstadials. The models simulate the exceptional character of the post-glacial warm interval only when the nonlinear terrestrial response to insolation forcing is included. Different rates of ice growth and decay, a nonlinear isostatic adjustment to the ice-load, thermal memory of the oceans, and so on, represent the terrestrial input^{6,7} (for a review, see ref. 8).

It has been argued, however, that the internal dynamics of the ice sheets can explain the dominant 100,000-yr periodicity of palaeoclimatic records^{9,10} and consequently, that the astronomical modulation of the Earth's climate has, at best, only a secondary role.

We have attempted to test whether the non-linear terrestrial response is necessary to explain the glacial-interglacial cycle. Is there an orbital configuration specific to an interglacial and can the exceptionally warm interglacial climate be explained by orbital configuration alone?

Interglacials are the key climatostratigraphic units recognized in Pleistocene strata worldwide—they reappear about every 10⁵ yr. It is the interglacial environment in which mankind flourished. Any useful prediction of future natural climates has to be able to differentiate between the interglacials and the glacials. If a specific orbital signature of an interglacial was found it would provide: (1) a better understanding of climate; (2) a definition of predictors of future gross climate shifts; and (3) improvement in the dating of past climates.

In this article we first summarize radiometrically dated evidence of climatic changes during the past 350,000 yr; then

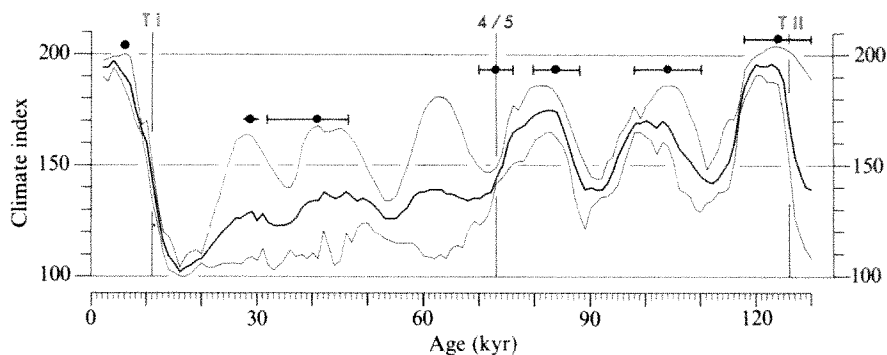
compare this record with the history of orbital perturbations and identify orbital configurations, if any, which correlate with interglacials, interstadials and stadials; and finally propose and test an index, based solely on orbital variables, with possible use as a predictor of past and future gross climate states.

Climatic variations of the past 350,000 yr

Past climates were reconstructed from indirect evidence on the size and volume of former ice sheets and on the change in habitats of flora and fauna^{11,12}. Radiometric dates, counts of annual layers and interpolated sedimentation rates provided the time scale. Ocean records are the most complete^{13,14}; land records are the most detailed and the best dated¹⁵. The accuracy of the dating decreases with time. Dating of the past 30,000 yr relies on radiocarbon; it is accurate to within $\sim \pm 1,000$ yr; in the preceding 10,000 yr to within about $\pm 5,000$ yr. Previous to that, because of contamination problems¹⁶, ¹⁴C ages have to be considered minimum values. U/Th isotopic ages give the highest accuracy ($\pm 5,000$ – $6,000$ yr) in the interval between 60,000 and 130,000 yr BP.

Figure 1 summarizes the climate records of the past 130,000 yr. It is based on three sets of proxy climate indicators dated by radiocarbon up to 50,000 yr BP and by U/Th decay series thereafter. The three groups are: (1) Pollen data from Grande Pile, France¹⁷⁻²⁰ and Tenaghi Phillipon, Greece^{21,22}, transferred onto a scale where 100 units correspond to the peak glacial (between 15,000 and 20,000 yr BP) and 200 units to the peak present interglacial (6,000 yr BP). (2) Sea level estimates from dated coral reefs in Barbados²³⁻³⁰, Curacao²⁴, Bermuda^{31,32}, New Guinea³³⁻³⁶, Japan³⁷, Loyalty Islands³⁸, Australia³⁹, Eniwetok Islands⁴⁰, Florida^{24,41} and the Bahamas^{41,42}, and ¹⁴C dated palaeo-sea levels in North Carolina^{43,44}, Scotland⁴⁵, and Australia⁴⁶, transferred on the same relative scale. (3) Delta ¹⁸O records^{3,47-49} in cores RC11-120,

Fig. 1 Climate severity index for the past 130,000 yr based on the three sets of proxy climate indicators (pollen, sea level and $\delta^{18}\text{O}$), expressed on a relative scale where 100 units correspond to the peak glacial and 200 to the peak of the present interglacial. Input data and procedure are explained in text. Horizontal bars show the dating accuracy whereas the envelope gives the range of the climate index values. Note large uncertainty between 20,000 and 70,000 yr BP.



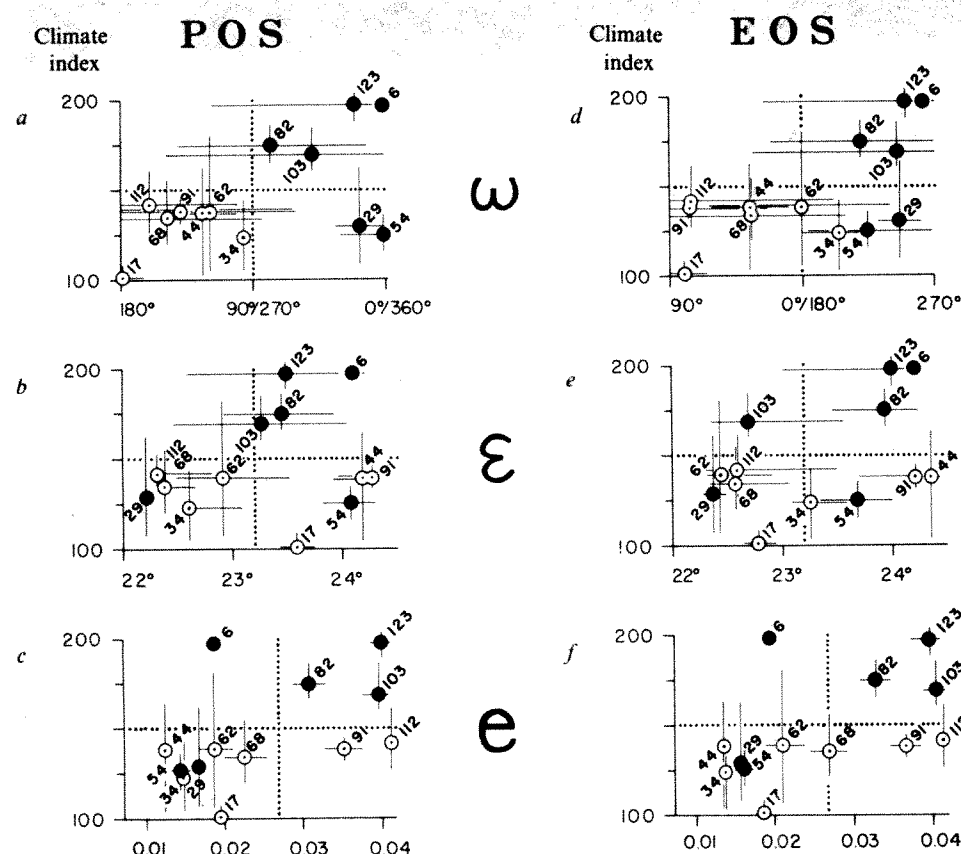


Fig. 2 Orbital signature at the peak (POS) and at the onset (EOS) of major warm and cold episodes of the past 130,000 yr, dated by radioisotopes. The episodes are identified by the age of their peak, given in thousands of years. Vertical scale: climate severity index from Fig. 1. Horizontal scales: longitude of the perihelion (ω) in degrees, obliquity (ϵ) in degrees, and the eccentricity (e). ● mark the episodes with the POS ω from 270° to 360° to 90°. Discussion in text.

METEOR 12392-1, V19-29, and V29-179 dated by correlation with sets 1 and 2 and by Toba Ash⁵⁰, and plotted along the same scale. The index was averaged arithmetically, first within each group and second among groups 1, 2 and 3. In the pollen record from Grande Pile¹⁷⁻²⁰ in France, and Tenaghi Philippon^{21,22} in Greece, the climate index γ was computed after the empirical formula:

$$\gamma = 4x + 2y + z$$

where x , y and z are the warm, cool, and cold principal pollen producing species, respectively, taken as a per cent of the total pollen count. Here x is the sum of oak, fir, hornbeam, alder and hazel; for Grande Pile yew and beech were added; for Tenaghi Philippon Mediterranean hornbeam and pistachio were added. y is the sum of pine, birch and juniper; with spruce added for France, and willow added for Greece. z is grass in France, and the sum of sagebrush and herbs in Greece.

The pollen formula was selected after multiple comparisons of the results with published palaeotemperature reconstructions of the past 130,000 yr (ref. 51). The above formula gave the closest agreement. Details of the procedures will be described elsewhere.

The smoothing process removed the relatively sharp boundaries observed in the land based records^{52,53} separating the cool and the temperate stages. The time series of most proxy climate indicators when inspected on a sufficiently detailed time scale, have a step function form⁵³. They show relatively steady intervals several millenia long, interspersed with short episodes of relatively rapid change. The more recent of these shifts were dated close to 10,000 yr BP (ref. 54), 23,000 yr BP (ref. 55), and 32,000 yr BP (ref. 55). Some such records may result from a step response of a proxy climate indicator with a broad climatic tolerance to a smooth continuous change of climate. However, recent data indicate that the climate over large regions changed in steps^{52,56} along a generally sinusoidal path.

The principal features of the palaeoclimatic history of the past 130,000 yr (shown in Fig. 1) can be summarized as follows. The past 10,000 yr and the interval between ~115,000 and 127,000 yr BP had an interglacial climate, considerably warmer than the rest of the time. There were only two interglacials in the

past 130,000 yr, both began by a pronounced stepwise environmental change accompanying a full scale deglaciation (termination). The most severe part of the last glacial lasted from ~22,000 to ~13,500 yr BP (ref. 57) and continental ice sheets were most extensive between 19,000 and 15,000 yr BP (ref. 58). At that time the sea level was 80 m (ref. 44), and possibly 130 m (ref. 43), lower than today. Similar environmental conditions and, by inference, climate, marked the next-to-last full glacial between 130,000 and 140,000 yr BP (ref. 15). Milder inter-

Table 1 Obliquity and June insolation at the start of the mild and cool intervals

Mild intervals			Cool intervals		
Time (kyr)	ϵ	June solstice insolation 90° N (Ly day ⁻¹)	Time (kyr)	ϵ	June solstice insolation 90° N (Ly day ⁻¹)
+10	22.61°	1,105	+19	23.16°	1,091
-11*	24.20°†	1,197	-1	23.58°	1,086
-34	22.60°	1,112	-22*	22.79°‡	1,048‡
-60	23.22°	1,146	-47	24.42°	1,132
-83	23.58°	1,197	-72	22.38°‡	1,018‡
-105	22.96°	1,185	-94	24.26°	1,075
-127*	24.04°†	1,239	-116	22.49°‡	991‡
-151	22.48°	1,132	-139*	23.57°	1,050
-176	23.75°	1,213	-164	23.87°	1,079
-198	22.85°	1,205	-187	22.62°‡	991‡
-220*	23.78°†	1,247	-209	24.29°	1,047‡
-242	23.25°	1,186	-231*	22.08°‡	972‡
-267	22.80°	1,129	-254	24.40°	1,104
-292	23.98°†	1,217	-280	22.90°	1,044‡
-314	22.54°	1,167	-303	23.52°	1,040‡
-335*	24.21°†	1,225	-324	23.36°	1,035‡
-355	22.36°	1,094	-345*	22.97°	1,049

* Warm interval following or cool interval preceding major deglaciation.

† Values equal or higher than that corresponding to a dated interglacial.

‡ Values equal or lower than that corresponding to the last peak glacial.

Table 2 ACLIN 1 cutoff values

Climate state	ACLIN 1 peak	Duration (kyr)
Interglacial	≥ 4.3	10
Temperate interstadial	$\geq 3.5 < 4.3$	8
Interstadial	$\geq 2.5 < 3.5$	4
Stadial and protostadial	< 2.5	

stadials punctuated the glacial between 115,000 and 22,000 yr BP: the first two episodes are sometimes referred to as temperate interstadials. The environment then locally approached, but globally did not reach, interglacial conditions. Remaining interstadials were cold with arctic climate in the northern middle latitudes. At $\sim 73,000$ yr BP (ref. 14) the climate deteriorated sharply.

As seen in Fig. 1, the climate severity between $\sim 70,000$ and 20,000 yr ago is relatively uncertain. While the deep-sea record and the pollen assemblages in Europe testify to harsh conditions and suggest a large volume of land based ice, the sea level estimates based on raised coral reefs imply a much smaller ice volume.

Palaeoclimates of the interval 130,000–350,000 yr BP are known in much lesser detail and are less accurately dated (Fig. 3). Four interglacials occurred in the past 350,000 yr; each was preceded by one of the major deglaciation dated to $\sim 11,000$, 127,000, 220,000 and 330,000 yr BP. The oldest date is based entirely on interpolated thicknesses of the Brunhes sediments in deep sea cores (ref. 15). Published estimates differ by about $\pm 10\%$. Continuous pollen bearing deposits⁵¹, soil-loess sequences¹⁵, and ^{18}O series⁵⁹ show that the frequency and amplitude of climate oscillations before and after 130,000 yr BP was approximately the same.

Climate index compared with orbital variables

The Earth circumsolar orbit is defined by three parameters:

- (1) The tilt ϵ of the Earth axis to the plane of the ecliptic (obliquity). It varies from $\sim 22^\circ$ to $\sim 24.5^\circ$ within a basic period of $\sim 41,000$ yr. With higher tilt, more insolation is directed to the high latitudes of both hemispheres in summer at the expense of low latitudes.
- (2) The longitude of the perihelion ω . This is measured as the angular distance of perihelion from the autumnal equinox, and defines the season in which the Earth makes its closest approach to the Sun. At perihelion, the intensity of solar radiation reaching the Earth is at a maximum. The precession of the equinoxes, as the process is called, has an irregular periodicity of 13,000–25,000 yr with periods of $\sim 23,000$ yr and 19,000 yr being the most common⁴.
- (3) The eccentricity e is the measure of the Sun's departure from the geometric centre of the Earth's orbit. High eccentricity increases the effect of ω which varies from close to 0 to ~ 0.06 , with periods of $\sim 100,000$ yr and 410,000 yr.

The three orbital parameters at any past or future time can be computed from the equations of planetary mechanics⁶⁸. Although we used recently revised calculations of Berger^{60,61}, the calculations of others agree to about 350,000 yr BP but they diverge further back in time⁶⁰. The precision of the calculations decreases with time due to the limited number of terms used in the expansions and to imprecise knowledge of the input parameters.

Figure 1 gives the timing of the major warm and cold climatic episodes with a wavelength of at least 10,000 yr. It also defines a starting (entry) date of each episode, half way between the neighbouring highs and lows. The orbital signature corresponding to the time of climate peaks, the peak orbital signature (POS), is plotted in Fig. 2a–c. Characteristics of the start of each episode, the entry orbital signature (EOS), is plotted in Fig. 2d–f. Each part of the figure is divided into quadrants defined by the mean values of the climate index and of the corresponding orbital measure.

Figure 2a shows, that mild or warm climatic episodes tend to peak when the longitude of perihelion ω is close to $0^\circ/360^\circ$ (September perihelion), whereas cool episodes culminate near 180° (March perihelion). We therefore label episodes with peak ω between 90° and 270° by open circles and the remainder by full circles. This convention applies throughout Fig. 2. Figure 2b shows that warm climates tend to be associated with high obliquity, and cool climates with low obliquity. Finally, (Fig. 2c) cold climates seem to be associated with low eccentricity.

None of the warm episodes display simultaneously high values of eccentricity, obliquity and autumn perihelion. For instance, the present interglacial peak at 6,000 yr BP has low eccentricity, the 29,000 yr BP interstadial peak has a low obliquity and eccentricity, and the relatively cold episode at 91,000 yr BP has a high obliquity and eccentricity, but a longitude of the perihelion close to 180° .

The orbital configuration disagrees only once with the corresponding climate, at 54,000 yr BP, a time of high obliquity and of autumn perihelion. Although the eccentricity is low, it is not much lower than at the peak of the present interglacial. Why, then, should 54,000 yr BP mark a cold climate? One possibility is that the disagreement stems from poor dating and possible incompleteness of the sea level curve between 40,000 and 65,000 yr BP, which strongly influenced the average index (Fig. 1). Contrary to the master curve, the 54,000-yr BP interval shows a relative amelioration in both the pollen data and the $\delta^{18}\text{O}$ record. However, as no marine terrace of such age is reported, we assume a low sea level between 45,000 and 60,000 yr BP.

Figure 2d–f shows the entry orbital signature as a function of the peak climate index. A phase shift of $\sim 5,000$ – $6,000$ yr exists between the plotted orbital variables and the climate index. The mild and warm episodes show ω close to 270° (June perihelion) whereas cool episodes close to 90° (December perihelion, Fig. 2d). Obliquity values are high during warm episodes and low during cold ones, with the same exceptions as in Fig. 2b. The eccentricity association is similar to Fig. 2c.

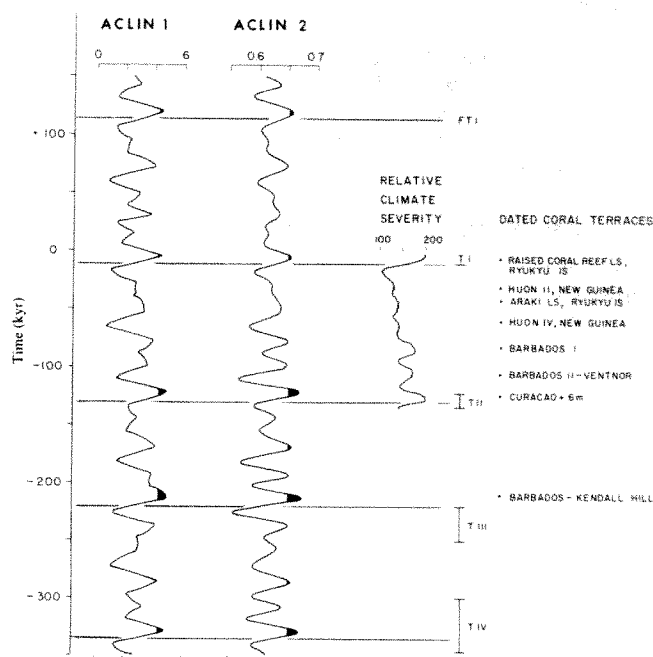


Fig. 3 Plots of ACLIN 1 and 2 for the past 350,000 yr and future 150,000 yr. Higher index value correlates with warmer climate. ACLIN 1 interglacial peaks ≥ 4.3 and ACLIN 2 peaks ≥ 0.654 filled in black. Climate severity index for the past 150,000 yr same as in Fig. 1. Mean ages of the U/Th dated coral reefs^{23,24,36,37} shown by dots. Vertical bars show the range of age estimates of terminations (TI–T IV) as obtained from deep sea cores¹⁵. FT I is the first expected future termination. Peaks as high or higher than that of 6,000 yr BP shown by arrow.

We conclude from Fig. 2 that the two interglacials, peaking at 6,000 yr BP and 123,000 yr BP, possess specific orbital characteristics which are absent during the remaining periods. They start at a time of June perihelion and reach a climax at a time of September perihelion, with high obliquity throughout the first half of the interval. Only the interstadial at 82,000 yr BP shows a similar, though markedly weaker character.

To test the correlation of high EOS obliquity with interglacials, all June perihelions ($\omega = 270^\circ$) during the past 360,000 yr BP were listed in Table 1, together with the corresponding obliquity. The known four interglacials started at 335,000, 220,000, 127,000 and 11,000 yr BP. The interglacial starting at 220,000 yr BP shows the lowest EOS obliquity of 23.8° , all three remaining interglacials start with higher values. All interstadials except the one at 292,000 yr BP have an obliquity lower than 23.8° . Thus, at least within the past 250,000 yr, an EOS obliquity of $\geq 23.8^\circ$ together with a June perihelion, is an exclusive characteristic of interglacials. However, this observation does not necessarily imply the existence of a physical mechanism directly linking the two observed properties with terrestrial climates. Some other variable indirectly related in space and time may provide the critical solar-terrestrial link.

It might be expected that the combined impact of high obliquity and of a June perihelion be fully reflected in the insolation values at high latitudes and that a single insolation curve for a selected latitude could serve as a predictor of interglacials. Table 1 shows that this is not the case. Even at the North Pole, where the departures are largest, the insolation at the onset of the postglacial was not any higher than during the four earlier interstadials. The physical explanation of the climatic importance of the high tilt probably lies in the dynamics of the seasonal sea-ice covers. These disintegrate in the late spring and summer as a result of high solar radiation^{62,63}. Only the obliquity cycle, but not the precessional cycle, increases the insolation of the pack ice fields of both hemispheres in phase.

The special importance of the autumn perihelion is linked to the dynamics of snow fields. The seasonal advance of snow fields in autumn has been closely correlated with the seasonal change of insolation⁶⁴, thus snow coverage may be more sensitive to insolation in autumn than during the rest of the year. If so, the autumn perihelion would delay the seasonal expansion of snow in the Northern Hemisphere but would have no measurable effect on the Southern Hemisphere, where seasonal snow fields are negligible.

The astronomical climate index

As shown in Fig. 2, high obliquity, high eccentricity, and autumn perihelion correlate with relatively warm climates. Based on this observation, we formulate an index which would predict a gross climate state as a function of ϵ , ω and e . The resulting astronomical climate index (ACLIN) should be high when ϵ and e is high and when ω approaches 360° . This approach leads to a purely empirical formula, ACLIN 1 (α_1), first introduced in 1978 (ref. 65):

$$\alpha_1 = \left| \frac{\omega_1 - 180}{90} \right| + \epsilon_2 - 22 + \frac{10^3 e_1^2}{2} \quad (1)$$

where α_1 is the climate index at time t_1 , and ω_1 , e_1 the longitude of the perihelion, and eccentricity at a time t_1 , and ϵ_2 the obliquity at a time t_2 , defined by:

$$\omega_2 = \omega_1 - 90 \text{ and } |t_1 - t_2| = \min$$

and where ω_2 is the longitude of the perihelion at a time t_2 . The separation of t_1 and t_2 by 90° in the longitude of the perihelion corresponds to a time lag of $\sim 5,000$ – $6,000$ yr implying a several millenia long memory in the climate system. Plots of α_1 are compared with palaeoclimatic evidence from the past 130,000 yr in Fig. 3, and for the past and future 1 Myr in Fig. 4.

While the correlation of ACLIN 1 with palaeoclimatic evidence is good, the separation of e and ω in two independent

variables disagrees with the laws of orbital geometry. An improved ACLIN formula therefore takes into account the physical link between orbital perturbation and insolation. It is composed of two terms: A , proportional to the relative intensity of ground level insolation at high latitudes at the summer solstice, and B , proportional to the relative intensity of insolation at the autumnal equinox.

$$A = \sin \epsilon \frac{1 + 2e \cos(\omega + 90)}{(1 - e^2)^2} \quad (2)$$

$$B = \frac{1 + 2e \cos \omega}{(1 - e^2)^2} \quad (3)$$

Both A and B are scanned at the same times t_1 and t_2 as in ACLIN 1. Other separations of t_1 and t_2 are possible but have not been investigated (see ref. 6). The present arbitrary selection of t_2 as used in all ACLIN formulas was based on the good correlation of EOS with the climate index. As the orbital variables are scanned at two different intervals, ACLIN is a time-lag bivariate model.

The basic ACLIN formula has the simplified form

$$\alpha_0 = a(A_1 + b \cdot B_1) + (A_2 + c \cdot B_2) \quad (4)$$

and can be written in full as

$$\alpha_0 = a \frac{\sin \epsilon_1 [1 + 2e_1 \cos(\omega_1 + 90)] + b(1 + 2e_1 \cos \omega_1)}{(1 - e_1^2)^2} + \frac{\sin \epsilon_2 [1 + 2e_2 \cos(\omega_2 + 90)] + c(1 + 2e_2 \cos \omega_2)}{(1 - e_2^2)^2} \quad (5)$$

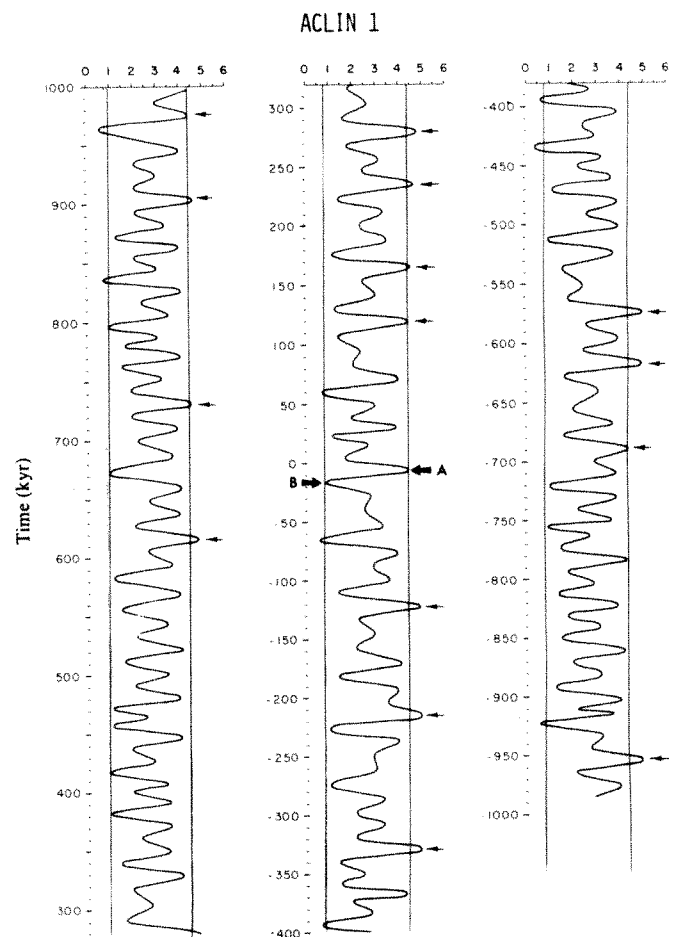


Fig. 4 Plot of ACLIN 1 from 1,000,000 yr BP to 1,000,000 yr AP. Higher value indicates warmer climate. Peak Holocene (A) level of ≥ 4.3 defines interglacials marked with arrows. Past peak glacial level at about 17,000 yr BP (B).

where α_0 is the ACLIN index for time t_1 , $e_1, \varepsilon_1, \omega_1$ the orbital elements at a time t_1 and $e_2, \varepsilon_2, \omega_2$ the orbital elements at a preceding interval t_2 defined in the same way as in ACLIN 1, and a, b and c the arbitrary coefficients.

Various versions of α_0 were studied. Best results were obtained with ACLIN 2 (α_2), in which $a = b = 0.25$ and $c = a^2$. A plot of ACLIN 2 is shown in Fig. 3.

The following criteria have been applied to test the predictive merits of ACLIN variants:

Accuracy level 1: predicts 10 interglacials in the past Myr, spread at $\sim 10^5$ yr intervals, it also fulfills criteria for level 3.

Accuracy level 2: predicts interglacials in the past 350,000 yr at times matching palaeoclimatic records within the accuracy of radiometric dating. Predicts temperate interstadials (and not interglacials) at $\sim 80,000$ and $100,000$ yr BP and the relative amplitudes of cold minima at $\sim 18,000, 65,000$ and $110,000$ yr BP.

Accuracy level 3: predicts the relative amplitude and timing of interglacials and temperate interstadials during the past 150,000 yr and the timing but not the relative amplitude of the cold peaks.

Accuracy level 4: predicts correctly the relative timing but not the amplitude of warm and cold intervals during the past 150,000 yr. Does not predict interglacials.

No solution has yet been found reaching accuracy level 1. ACLIN 1 is the only formula tested so far which reached accuracy level 2. Formula ACLIN 2 reached accuracy level 3. Additional ACLIN variants can be constructed to obtain a better agreement with palaeoclimatic records, following the tuning strategy proposed by Imbrie and Imbrie⁸. However, the two described versions are sufficiently successful to be used for long range climate estimates.

ACLIN time series approach a smooth sinusoidal function. A computer program can transform ACLIN into a stepwise form, scanning for maxima and minima (Table 2), and assigning a parametrized duration to the three mild climate states. In such a model, protostadials and stadials occupy the rest of the time. Protostadials are gross cold climate episodes characterized by a restricted size of continental ice sheets which occur within the first 50,000 yr or so of an interglacial-glacial cycle.

Figure 5 shows the spectral response of ACLIN 1 and ACLIN 2 compared with amplitude spectra of the deep sea core V28-238 and the spectrum of Imbrie and Imbrie's model. ACLIN 2 has significant peaks only at periods of $\sim 41,000$ yr (the obliquity cycle) and $\sim 23,000$ and $19,000$ yr (the precession cycle) whereas ACLIN 1 shows an additional significant peak near $100,000$ yr. Although the $100,000$ -yr peak in ACLIN 1 is not as dominant as it is in the oxygen isotope record of core V28-238, we find the ACLIN 1 result encouraging. It shows that a bivariate model based solely on the orbital input is capable of generating a $100,000$ -yr cycle. In ACLIN this period is due to the eccentricity, but the real cause is the coupling of high obliquity and precession modulated by the eccentricity (see Table 1).

The spectrum of ACLIN 1, based entirely on orbital variables, closely resembles the spectrum of Imbrie and Imbrie's model⁸ in which nonlinear ice growth rates were combined with an insolation function. Similar to Imbrie's model, ACLIN 1 has a power in the $400,000$ -yr range but with a lower and more realistic peak.

Discussion

Given the complexity of the Earth-atmosphere system, it is surprising that the simple ACLIN formulae successfully simulate the principal features of the palaeoclimatic record. This emphasizes again the dominant role of orbital variations in the shaping of past climates. The most significant achievement of ACLIN is the successful discrimination of interglacials on the sole basis of orbital input scanned across a $\sim 5,000$ -yr interval. All earlier models had to include a nonlinear response of the terrestrial system to orbital forcing to separate the interglacials and generate the $100,000$ -yr periodicity.

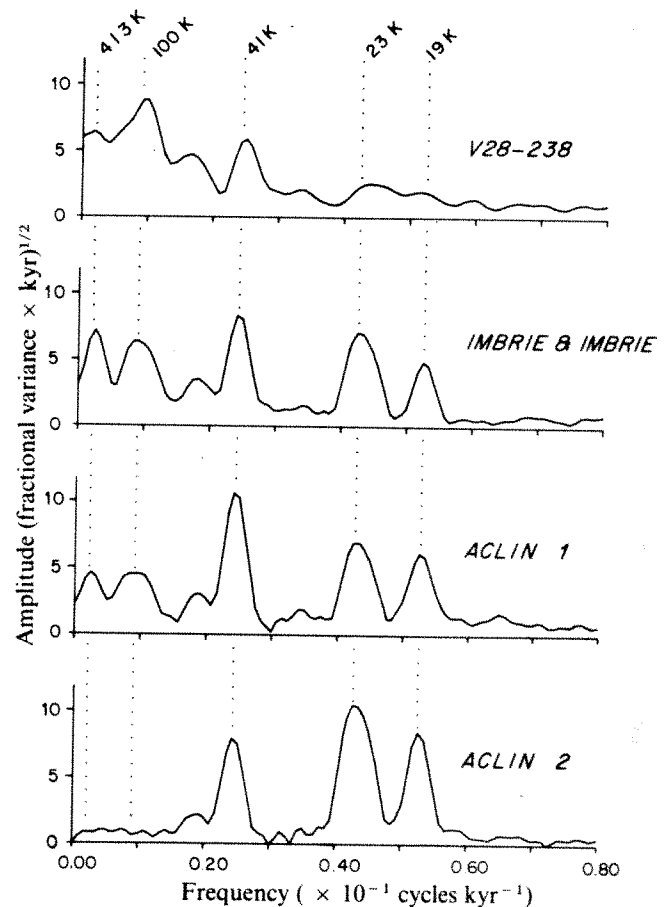


Fig. 5 Power spectra of the oxygen isotope record in the Pacific core V28-238 over the past 730,000 yr (ref. 59); the output of the Imbrie and Imbrie non-linear climate model⁸ for the past 1,000,000 yr; the ACLIN 1 model over the past 1,000,000 yr; and the ACLIN 2 model over the same interval. Observe a $\sim 100,000$ and $\sim 410,000$ yr peak present in the ACLIN-1 and in the Imbrie model, but missing in ACLIN 2.

While utmost caution is needed in the interpretation of our data⁸, the results have some implications for climate theory:

- (1) The gross climate state may almost be in equilibrium with orbital forcing. Otherwise, it would seem impossible to predict interglacials solely from the orbital configuration of a $\sim 5,000$ yr long interval.
- (2) The primary origin of the $100,000$ yr cycle can be orbital rather than terrestrial. If this is so, a primary orbital forcing would be modulated by a nonlinear thermal, mass, and isostatic response of ice and oceans, rather than the reverse.
- (3) Increased insolation to high latitudes at the expense of low latitudes is associated with warm global climates, not the other way around.
- (4) The interaction of both hemispheres, not the sole dominance of the Northern Hemisphere, seems to be critical in the dynamics of the $100,000$ yr cycle. This is indicated by the dominant role of obliquity in the ACLIN formula. (It had been suggested earlier^{1,2} that the tilt may have played a principal role in shaping the past climates.)
- (5) The seasonal pack-ice in both hemispheres and the snow fields in the Northern Hemisphere seem to play a key part in the dynamics of Pleistocene climates as suggested earlier^{62,66}.

While the ACLIN simulates the relative intensity of mild events quite well, it does not reproduce the cold episodes very well. For instance, the $133,000$ -yr BP ACLIN 1 value is considerably higher than that for $17,000$ yr BP, while the extent and the volume of continental ice was approximately the same at both intervals. Interglacials indicated by proxy data at $\sim 400,000$ and $500,000$ yr BP, are also poorly reproduced by ACLIN. ACLIN peaks at corresponding times do not reach the

4.3 cutoff level. However, the accuracy of both geological and astronomical data before 350,000 yr BP is low.

The general agreement of the ACLIN time series with the reconstructed palaeoclimates permits future gross climate states to be predicted. Such an estimate, however, can only be valid if man's impact on the land and the atmosphere has not significantly modified the mechanism of climate change, and will not do so in the future.

The ACLIN time series is extrapolated into the future in Figs 3 and 4. Both α_1 and α_2 predict a considerably cooler climate for the next 114,000 yr, followed by an interglacial. The first cold peak with a period of 10,000–100,000 yr should occur 4,000 yr AP (after present). ACLIN 1, but not ACLIN 2, predicts an amelioration around 15,000 yr AP followed by a cold interval centred around 23,000 yr AP. This agrees with the

prediction of Berger *et al.*⁶⁷ and Imbrie and Imbrie⁸. A major glaciation, comparable to stage 4 of the last glacial cycle, is indicated at 60,000 yr AP.

Our approach can predict at best only a very approximate general framework of gross climate changes. Detailed climate reconstructions or predictions will have to consider interactions of orbital perturbations with changing solar activity and with natural and manmade variables of terrestrial origin.

We thank J. Imbrie for innumerable improvements as well as for Fig. 5 data, and N. Shackleton, K. Bergman, J. Hays, and W. S. Broecker for reading the manuscript. Research was supported by NSF grants ATM77-28522 and ATM80-01470, and the Office of Nuclear Waste Isolation, managed by Battelle Memorial Institute for the Department of Energy under contract DE-AC06-76-RLO-1830.

Received 27 May; accepted 8 December 1980.

1. Emiliani, G. & Geiss, J. *Geol. Rdsch.* **46**, 576–601 (1959).
2. Broecker, W. S. *Science* **151**, 299–304 (1966).
3. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
4. Berger, A. *Nature* **269**, 44–45 (1977).
5. Goreau, T. J. *Nature* **287**, 620–622 (1980).
6. Calder, N. *Nature* **252**, 216–218 (1974).
7. Chin, W. Q. & Yevjevich, V. *Hydrology Papers* 65 (Colorado State University, 1974).
8. Imbrie, J. & Imbrie, J. Z. *Science* **207**, 943–953 (1980).
9. Oerlemans, J. *Nature* **287**, 436–438 (1980).
10. Mitchell, J. M. Jr in *The Quaternary of the United States, INQUA VII Cong.* (eds Wright, H. E. Jr & Frey, D. G.) 381–901 (Princeton University Press, 1965).
11. Flint, R. F. *Glacial and Quaternary Geology* (Wiley, New York, 1971).
12. Woldstedt, P. *Das Eiszeitalter* 1, 2, 3 (Enke, Stuttgart, 1954).
13. Emiliani, C. J. *Geol.* **74**, 109–124 (1966).
14. Ruddiman, W. F. & McIntyre, A. *Geol. Soc. Am. Mem.* **145**, 111–146 (1976).
15. Kukla, G. *Earth Sci. Rev.* **13**, 307–374 (1977).
16. Grootes, P. M. *Science* **200**, 11–15 (1978).
17. Woillard, G. in *IGCP Proj. 73/1/24, Rep. 6* (ed. Sibrava, V.) (CSSR, 1980).
18. Woillard, G. *Acta geogr. slov.* **14**, 1–118 (1975).
19. Woillard, G. *Quat. Res.* **9**, 1–21 (1978).
20. Woillard, G. *Bull. Soc. belge Géol.* **88**, 51–69 (1979).
21. Wijmstra, T. A. *Acta bot. neerl.* **4**, 511–527 (1969).
22. Wijmstra, T. A., & Smit, A. *Acta bot. neerl.* **25**, 297–312 (1976).
23. Broecker, W. S. *et al. Science* **159**, 297–300 (1968).
24. Harmon, R. S., Ku, T.-L., Matthews, R. K. & Smart, P. L. *Geology* **7**, 405–409 (1979).
25. James, N., Mountjoy, E. & Omura, A. *Bull. geol. Soc. Am.* **82**, 2011–2018 (1971).
26. Steinen, R. P., Harrison, R. S. & Matthews, R. K. *Bull. geol. Soc. Am.* **84**, 63–70 (1973).
27. Mesolella, K. T., Matthews, R. K., Broecker, W. S. & Thurber, D. L. *J. Geol.* **77**, 250–274 (1969).
28. Fairbanks, R. K. & Matthews, R. K. *Quat. Res.* **10**, 181–196 (1978).
29. Matthews, R. K. *Quat. Res.* **3**, 147–153 (1973).
30. Bender, M. L. *et al. Bull. geol. Soc. Am. Part I*, **90**, 577–594 (1979).
31. Land, L. S., MacKenzie, F. T. & Gould, S. J. *Bull. geol. Soc. Am.* **78**, 993–1006 (1967).
32. Hermon, R. S., Schwarcz, H. P. & Ford, D. C. *Quat. Res.* **9**, 205–218 (1978).
33. Veeh, H. H. & Chappell, J. M. A. *Science* **167**, 862–865 (1970).
34. Chappell, J. & Veeh, H. H. *Nature* **276**, 602–604 (1978).
35. Chappell, J. *Bull. geol. Soc. Am.* **85**, 553–570 (1974).
36. Bloom, A. L., Broecker, W. S., Chappell, J. M. A., Matthews, R. K. & Mesolella, K. J. *Quat. Res.* **4**, 185–205 (1974).
37. Konishi, K., Schlanger, S. & Omura, A. *Mar. Geol.* **9**, 225–240 (1970).
38. Marshall, J. F. & Launary, J. *Quat. Res.* **9**, 186–192 (1978).
39. Veeh, H. H., Schwebel, D., Van de Graaf, W. J. E. & Denman, P. D. J. *geol. Soc. Austr.* **26**, 285–292 (1979).
40. Thurber, D., Broecker, W., Blanchard, R. & Potratz, H. *Science* **149**, 55–58 (1965).
41. Broecker, W. & Thurber, D. *Science* **149**, 58–60 (1965).
42. Neumann, A. C., & Moore, W. S. *Quat. Res.* **5**, 215–224 (1975).
43. Emery, K. O. & Merrill, A. S. *Bull. geol. Soc. Am.* **90**, 689–694 (1979).
44. MacIntyre, I. G., Pilkey, O. H. & Stuckenrath, R. *Bull. geol. Soc. Am.* **90**, 692–694 (1979).
45. Cullingford, R. A., Caseldine, C. J. & Gotts, P. E. *Nature* **284**, 159–161 (1980).
46. Jones, B. G., Young, R. W., & Eliot, I. G. J. *geol. Soc. Austr.* **26**, 255–264 (1979).
47. Shackleton, N. J. in *The Fate of Fossil Fuel* (eds Anderson, N. & Malahoff, A.) 401–427 (Plenum, New York, 1977).
48. Ninkovich, D. & Shackleton, N. J. *Earth planet. Sci. Lett.* **27**, 20–34 (1975).
49. Streeter, S. S. & Shackleton, N. J. *Science* **203**, 168–171 (1979).
50. Ninkovich, D., Shackleton, N. J., Abdel-Monem, A. A., Obradovich, J. D. & Izett, G. *Nature* **276**, 574–577 (1978).
51. Van Der Hammen, T., Wijmstra, T. A. & Zagwijn, W. H. in *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 391–424 (Yale University Press, New Haven, 1971).
52. Woillard, G. *Nature* **281**, 558–562 (1979).
53. Bryson, R. A. & Wendland, W. M. in *Life, Land and Water* (ed. Mayer-Oakes, W. J.) 271–298 (University of Manitoba Press, Winnipeg, 1967).
54. Broecker, W. S., Ewing, M. & Heezen, B. C. *Am. J. Sci.* **258**, 429–448 (1960).
55. Morner, N.-A. *Can. J. Earth Sci.* **8**, 1423–1431 (1971).
56. Kukla, G. in *Paleoecology of Africa* (ed. van Zinderen, E. M.) (Bakker, The Netherlands, in the press).
57. Morner, N.-A. & Dreimanis, A. *Geol. Soc. Am. Mem.* **136**, 107–134 (1973).
58. Dreimanis, A. & Goldthwait, R. P. *Geol. Soc. Am. Mem.* **136**, 71–106 (1973).
59. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39 (1973).
60. Berger, A. *Astr. Astrophys.* **51**, 127–135 (1976).
61. Berger, A. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
62. Kukla, G. J. in *Climate Change* (ed. Gribbin, J.) 114–129. (Cambridge University Press, 1978).
63. Washington, W. M. & Chervin, R. M. *J. Appl. Met.* **18**, 3–16 (1979).
64. Kukla, G. J. *Nature* **253**, 600–603 (1975).
65. Kukla, G. in *Probability of expected climate stresses in North America in the next one million years: Topical Report* (Battelle, Pacific Northwest Laboratories, 1978).
66. Croll, J. *Climate and Time in the Geological Relations: A Theory of Secular Changes of the Earth's Climate* 4th edn (London, 1975).
67. Berger, A., Guioit, J. & Kukla, G. in *Proc. int. Scientific Assembly, Serbian Acad. Sci. Arts, 1979* (in the press).
68. Milankovitch, M. M. *Beograd. Koninglich-Serbische Akademie* 484 (English transl. by Israel Program for Scientific Translation and published for the U.S. Department of Commerce and the National Science Foundation, Washington DC, 1941).

A carbon budget for overfishing off Peru

John J. Walsh

Oceanographic Sciences Division, Brookhaven National Laboratory, Upton, New York 11973, USA

The rise and collapse of the Peruvian anchovy fishery between the 1950s and the 1970s was accompanied by marked changes in the fluxes of a carbon budget for the upwelling ecosystem.

THE anchovy harvest off the coast of Peru increased from 700,000 tonnes yr⁻¹ in 1958 to a maximum of ~12 million tonnes yr⁻¹ in 1970, but then declined precipitously to ~2 million tonnes yr⁻¹ in 1973 and to ~1 million tonnes yr⁻¹ between 1977 and 1979. The collapse of the anchovy fishery, located in the upwelling ecosystem off the coast of Peru, is attributable to overfishing and the diminished survival of anchovy larvae during warm-water intrusions (El Niños) southwards along the coast of Peru in both 1972 and 1976¹.

The very large yield of anchovy (*Engraulis ringens*) in the heyday of the fishery is usually attributed to high primary production in the upwelling system and the high efficiency of the

food chain in which the commercial fish feed directly on phytoplankton. In reality, there has been debate^{2,3} for more than a decade about the relative importance in the maintenance of the anchovy stocks of the short food web (phytoplankton → anchovy) with ~10% ecological efficiency and of a longer food web (phytoplankton → zooplankton → anchovy) with efficiency of ~1% but functioning over an area perhaps 10 times as great. A primitive simulation⁴ of the daily nitrogen flux through the Peruvian upwelling system suggested that the longer and more widely spread food-web might be more important during the strong winter upwelling off Peru. A clearer understanding of the trophodynamics of the upwelling system is now possible by

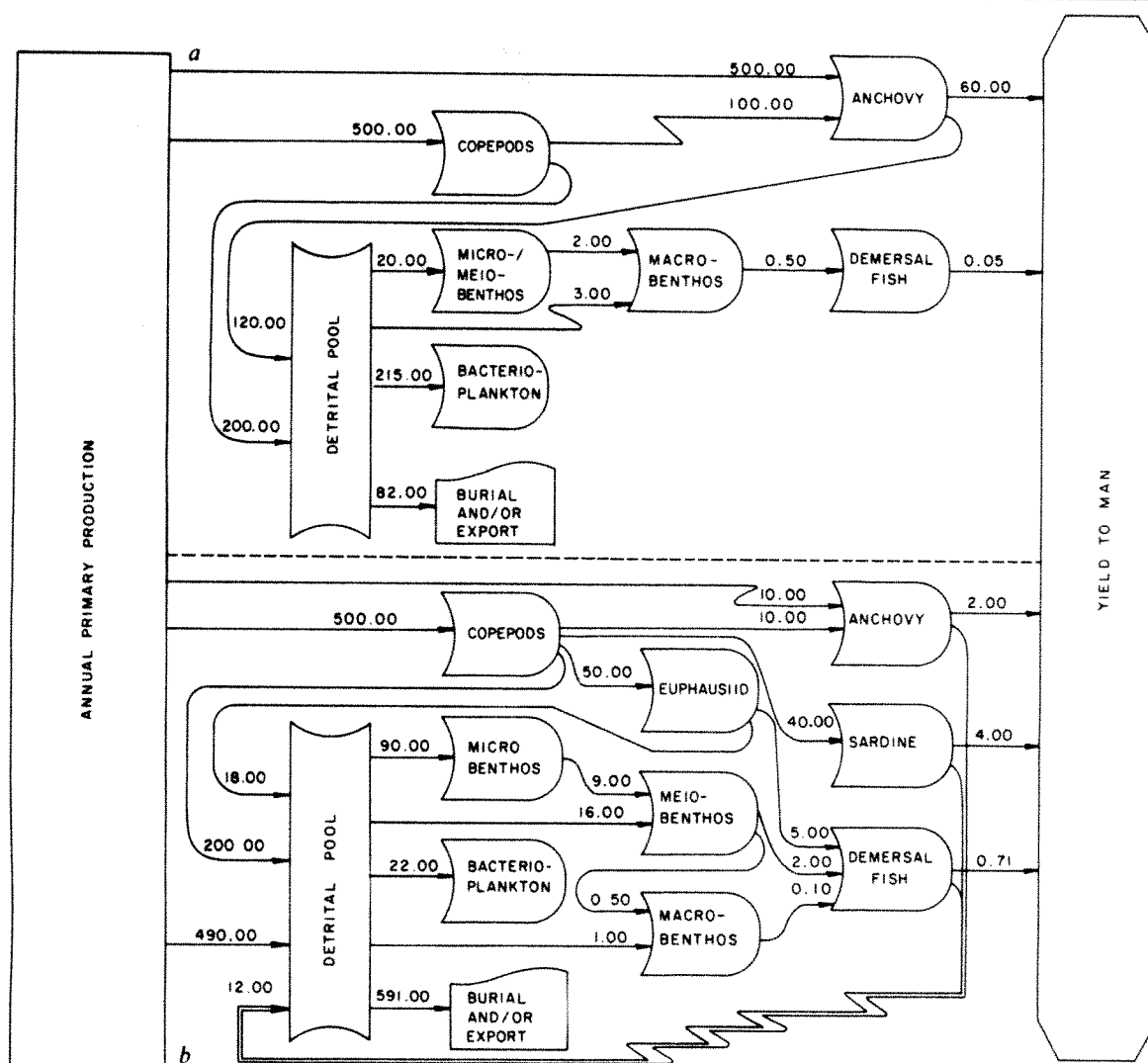


Fig. 1 A budget for the flux of carbon ($\text{g C m}^{-2} \text{yr}^{-1}$) through the Peru food web: a, before (~1966-69) and b, after (~1976-79) overfishing of anchovy.

examining the annual carbon budgets before and after overfishing, in the periods 1966-69 and 1976-79 respectively. The 1966 primary production data (Table 1) are high because they were collected while the ship was drifting within a high chlorophyll area at the shelf-break⁴. The March-April 1976-77 data are low because repeated time series stations were taken within a low chlorophyll area near the coast¹. Other data in Table 1 were averaged from transects taken across the shelf. High chlorophyll and abundant anchovy eggs are found within 50 km of the coast¹, suggesting that a combined productivity average of this region is a reasonable estimate of the yearly carbon input to the Peru upwelling food chain. At a rate of at least $3 \text{ g C m}^{-2} \text{day}^{-1}$ ($\bar{X} = 4 \text{ g C m}^{-2} \text{day}^{-1}$ in Table 1), the annual primary production is a minimum of $\sim 1,000 \text{ g C m}^{-2} \text{yr}^{-1}$ off Peru (see Table 2) in contrast to $\sim 200\text{--}500 \text{ g C m}^{-2} \text{yr}^{-1}$ for shelf ecosystems at higher latitudes⁵.

An upwelling zone of $\sim 50 \text{ km}$ offshore width, extending along the Peru-Chile coast of $\sim 2,000 \text{ km}$ length ($4\text{--}22^\circ\text{S}$), and a peak fish catch of 10 million tonnes yr^{-1} harvested over the whole area ($1 \times 10^5 \text{ km}^2$) suggests a yield of at least 100 tonnes $\text{km}^{-2} \text{yr}^{-1}$. Such a harvest was an order of magnitude higher than that presently derived from the North Sea, the Mid-Atlantic Bight or the Bering Sea⁵. With a conversion factor of $0.06 \text{ g C} = 1 \text{ g ww}$, a harvest of 100 tonnes $\text{km}^{-2} \text{yr}^{-1}$ ($100 \text{ g m}^{-2} \text{yr}^{-1}$) suggests, however, a yield to man of only $6 \text{ g C m}^{-2} \text{yr}^{-1}$ of anchovy off Peru from a primary production of at least $\sim 1,000 \text{ g C m}^{-2} \text{yr}^{-1}$. As man was the major consumer of the anchovy during its peak and declining yields, this harvest may be a reasonable estimate of fish secondary production over the

whole region. It suggests a food chain efficiency of $< 1\%$ before the fishery collapsed. The yield from the anchovy fishery in 1976-79 was ~ 1 million tonnes yr^{-1} , or an implied food chain efficiency to man of only 0.1% after overfishing if the catch then also approximated secondary production of the anchovy.

The same maximum estimate of secondary production, taken from a 90% smaller upwelling region, would represent a larger yield to man per unit area—a greater implied food chain efficiency. For example, if most of the peak yield had been derived only from the area of high anchovy egg and larvae abundance¹, north of Callao on the broad shelf between 6 and 10°S , then the annual harvest would have been $60 \text{ g C m}^{-2} \text{yr}^{-1}$ of fish before overfishing occurred (Table 2). If the anchovy in this area also consumed phytoplankton half the year and zooplankton with a growth efficiency of 20% (Table 2) during the other half of the year⁶, the fish yield would then have been 10% of their food input from this smaller region between 6 and 10°S . A more detailed carbon budget for the food web off Peru may determine the fate of the primary production along 90% of the Peru coast before overfishing, if most of the annual yield was derived from facultative herbivorous fish within an area of only $1 \times 10^4 \text{ km}^2$.

Carbon budget off Peru

Information on oxygen content, particle fluxes to the bottom, sea-bed respiration, organic content of the sediments and nutrient depletion of the water column aid in the construction of a more complete carbon budget before (Fig. 1a) and after overfishing (Fig. 1b). Previous analysis of the demersal part of

the nearshore food chain before overfishing suggested that the input of carbon to the bottom consisted mainly of faecal pellets rather than phytodetritus⁴. In terms of carbon consumption, the decomposition of the faecal pellets could be attributed to caprophagy and/or bacterial activity in both the water column and sediments. Additional calculations⁷ had indicated that the oxygen demand of this detrital carbon flux was such that most of the faecal pellets could be remineralized within the upper 60 m of the water column. After overfishing and the decline of grazing pressure from adult anchovy, however, a large flux of phyto-detritus may have sunk out of the water column when nutrient depletion occurred at the edge of the shelf.

The estimate for bacterioplankton production (Table 2) from consumption of faecal pellets within the upper part of the water column before overfishing (Fig. 1a) was obtained from a steady-state budget of the vertical decline of oxygen at 15°S in 1966–69; 4.5 ml O₂ l⁻¹ above 20 m and 0.5 ml O₂ l⁻¹ below⁷. As a result of oxygen depletion in upwelled source water during March to April 1976, 1977 and 1978, however, the average O₂ concentration at 15°S was ~1 ml O₂ l⁻¹ in the upper 20 and 0.5 ml O₂ l⁻¹ below, that is a vertical gradient of ~0.04 ml O₂ l⁻¹ m⁻¹ after overfishing in contrast to 0.4 ml O₂ l⁻¹ m⁻¹ a decade earlier. The local production of aerobic bacteria was thus estimated to be an order of magnitude less (Table 2) in the low chlorophyllic surface waters of the second carbon budget (Fig. 1b) with the remaining detrital carbon then sinking to where more denitrification⁸ occurred in deeper water and sulphate reduction⁹ in the bottom sediments.

Table 1 Mean daily primary production in Peru Shelf waters from 1966 to 1978

Year	g C m ⁻² day ⁻¹	(N)	Month	Area	Ref.
1965	6.3	(15)	March–April	15°S	35
1969	5.2	(32)	March–April	15°S	35
1969	1.3	(12)	June	10–15°S	4
1974	4.5	(5)	February–March	7–8°S	19
1975	3.8	(5)	March–May	6–12°S	36
1976	3.2	(4)	August–September	10–15°S	1
1976	1.7	(29)	March–April	15°S	35
1977	1.9	(43)	March–April	15°S	35
1977	7.3	(10)	November	9–10°S	18
1978	4.3	(30)	March–April	15°S	35

$$\bar{X} = 3.95.$$

The aerobic carbon production of the benthos before overfishing (Fig. 1a) was estimated from both P/B estimates and bottom respiration measurements at 15°S in March to April 1969¹⁰. The mean macrobenthos biomass across the shelf in 1969¹¹ was 0.5 g C m⁻², suggesting a secondary production of 0.5 g C m⁻² yr⁻¹ (Table 2) and an ingestion demand of 5.0 g C m⁻² yr⁻¹ from a P/B ratio of 1 and a growth efficiency of 10%. Approximately 20% of the measured ~25 g C m⁻² yr⁻¹ bottom respiration¹⁰ might thus be due to macrobenthos off Peru in 1969. The rest is attributed to micro- and meio-benthos as in the bottom metabolism of other shelves^{12,13}.

The total input of carbon to the shelf bottom before overfishing (Table 2) is similar to a flux of 88 g C m⁻² yr⁻¹ at 50 m measured with floating sediment traps on the narrow shelf at 15°S in 1977⁹. If a higher sinking loss was applied (without other losses) to an early model of the nearshore Peru upwelling system¹⁴, the predicted amount of downstream chlorophyll in the offshore region was less than that observed. The nearshore area at 15°S contained the same amount of low chlorophyll (<2 µg l⁻¹) during 1966, 1969, 1976, 1977 and 1978, as that found inshore at 10°S during 1960–69 and 1976–77, suggesting that an estimate of a sinking loss of only 10% of nearshore primary production is reasonable, as higher phytoplankton biomass is found downstream, that is, offshore.

The production of the benthic food web after overfishing (Fig. 1b) was first estimated from sulphate reduction estimates of 90 g C m⁻² yr⁻¹ for anaerobic metabolic demand at 15°S in

surface sediments⁹. The sulphur bacterium, *Thioploca*, was found on the bottom in large numbers along the whole Peru coast during 1976, and the near bottom oxygen concentration at 15°S on the 100-m isobath was <0.1 ml O₂ l⁻¹ in 1976–77, compared with ~0.2–1.0 ml O₂ l⁻¹ in 1969. The content of the secondary nitrite maximum increased from 3–6 µg-atom NO₂ l⁻¹ in 1966–69¹⁵ to 14 µg-atom NO₂ l⁻¹ in 1977⁸. Similarly no sulphate reduction was detected in the water column in 1969¹⁶ at 15°S, whereas H₂S was found there in March to April 1976¹⁷ and at 8–9°S during November 1977¹⁸. Based on an average 1977 meio-benthic biomass of 0.6 g C m⁻² across the shelf⁹ and a P/B ratio of 4/1 reflecting a higher turnover rate of these small organisms, the secondary production of the meio-benthos might have been 2.5 g C m⁻² yr⁻¹ (Table 2), 28% of the sulphate reduction estimate of microbial production. Quantitative macrobenthic data are not available for 1976–78 but their biomass and production (Table 2) were then presumably smaller because more extensive anoxic conditions occurred on the Peru shelf.

As the primary electron acceptor in the oxidation of organic matter in the sea is oxygen, then nitrate in the process of denitrification, and finally sulphate reduction, the above sequence suggests that a larger than normal organic loading to the bottom of the shelf and upper slope had occurred by 1976–77 (Fig. 1b). The total detrital flux from the water column after overfishing in 1976–79 is, in fact, estimated (Table 2) to be 70% of the 1,000 g C m⁻² yr⁻¹ primary production. After the 1976 El Niño, such a detrital loss (Fig. 1b) is almost an order of magnitude larger than that estimated for 1966–69 (Fig. 1a), but similar to an estimate of 64% detrital loss¹⁹ in 1974 at 7–8°S after the 1972 El Niño. Within the geological record, a 10-fold increase in sedimentation rate is associated with a doubling of the organic carbon content of the sediments²⁰. If the above carbon budgets are correct, a similar process may have occurred off Peru from 1966–69 to 1976–79.

The distribution of organic carbon in the surface sediments (Fig. 2) suggests that before overfishing in 1968 and 1969^{11,21,22} the sediment carbon off Peru increased both seawards at the slope and polewards. Without consideration of changes in food chain dynamics, a primary production rate of 1,000 g C m⁻² yr⁻¹ and a previous sedimentation rate of 0.1 cm yr⁻¹ can be calculated²⁰ as a carbon accumulation rate within the sediments of 114 g C m⁻² yr⁻¹ off Peru, that is similar to that from the carbon budget before overfishing (Table 2). The location of the 2% carbon isopleth in the sediments of the wide shelf north of Callao in 1968–69 (Fig. 2) also suggests that less carbon used to be lost from the water column at 10°S than 15°S. These spatial trends continued after overfishing in 1976 (C. Delgado, L. A. Flores and G. T. Rowe, unpublished data) and 1977^{18,23}, but the carbon content of the surface sediments may have doubled recently. The same primary production and a sixfold increase in sedimentation rate (Table 2) lead to an empirical carbon accumulation rate²⁰ of 684 g C m⁻² yr⁻¹ in the sediments compared with that of 591 g C m⁻² yr⁻¹ estimated from the overfishing budget (Fig. 1b).

At 15°S, 4.5% carbon was the highest surface sediment value observed on the shelf in March to April 1969¹¹, but 10% carbon was found at the same stations with the same methods in 1976. Furthermore, within 22 cores taken across the 15°S Peru shelf and slope in 1977–78, the mean subsurface C/N value at 50 cm depth of these cores was 7.76 compared with 7.53 in the surface sediments, suggesting little diagenetic change²⁴. However, the mean % carbon was 2.89 at 50 cm depth and 4.82 at the surface, implying a recent increase in carbon flux to the sediments that could have occurred within the past decade. Inadequate spatial sampling means that areas of high organic content may have been missed in these 1968–78 surveys, but the horizontal, temporal and vertical trends of the sediment carbon data are at least consistent with the carbon budgets of the water column. The first carbon budget (Fig. 1a) for the food web before overfishing may be most applicable to northern Peru waters during 1966–69. The second budget (Fig. 1b) may instead reflect

the food web both before overfishing off southern Peru and afterwards off northern Peru.

The predicted change in dominance of pelagic and demersal fish species after overfishing (Fig. 1b) reflects the decline in these budgets of anchovy ingestion demands on both phytoplankton and zooplankton⁶. The increase of zooplankton biomass after the 1972 and 1976 El Niños was mainly in the standing stocks of omnivorous euphausiids^{1,19}. In contrast, there was little apparent change in their biomass after the 1957 El Niño²⁵, when the anchovy fishery was just beginning. The sardine, *Sardinops sagax*, and hake, *Merluccius gayi*, both eat zooplankton off Peru, suggesting that the stocks of these fish should also increase after the anchovy decline. Similar changes occurred in demersal species abundance within the North Sea after the pelagic fish stocks were overfished²⁶.

Following the 1976 El Niño, sardine indeed constituted ~67% of the pelagic yield of 1.5 million tonnes yr⁻¹ in 1977. During 1978 and 1979, the catch of mackerel, sardine, pipefish, and horse-mackerel was still ~50% of the total pelagic yield of 1–2 million tonnes yr⁻¹. Furthermore, the potential demersal yield of 'virgin' hake stocks was estimated to be 0.75 million tonnes yr⁻¹ after the 1972 El Niño²⁷. This is similar to the change predicted in the second carbon budget (Fig. 1b) of a yield of 0.71 g C m⁻² yr⁻¹, or 1.2 million tonnes yr⁻¹ of demersal fish along the whole coast (1 × 10⁵ km²). Despite restrictions in fishing, however, the same amount of plankton was not consumed by euphausiids, sardine, mackerel and hake in 1976 as by anchovy in 1966. Instead, the oxygen content of the euphotic zone and sea bed has since declined, the nitrite and sulphide content of the water column increased, and the organic content of the bottom sediments also apparently increased after overfishing.

Nitrogen mass balance for carbon fluxes

If 10³ tonnes km⁻² yr⁻¹ of anchovy had been produced on the wide shelf north of Callao (1 × 10⁴ km²) and 10² tonnes km⁻² yr⁻¹ along the rest of the coast (9 × 10⁴ km²), the combined secondary production of 19 million tonnes yr⁻¹ before overfishing is the same as that originally estimated in 1969^{2,3}. The carbon budget for the Peru food chain after overfishing (Fig. 1b) is thus a more reasonable description of the food chain along most of the Peru coast before and after overfishing with the exception of the 6–10°S area. If this second budget did apply to 15°S during 1966–69 (before overfishing), then about half of the daily production of phytoplankton was always lost as sinking detritus to the upper slope after nutrient depletion within offshore waters²⁸ rather than cropped inshore by zooplankton and anchovy^{4,29}. Annual carbon budgets of the shelf food webs of the Mid-Atlantic Bight⁵, the Gulf of Mexico and the Bering Sea similarly suggest that ~50% of their primary production

Table 2 Changes in annual carbon fluxes of the Peru upwelling ecosystem between 1966–69 and 1976–79

	1966–69 (Ref.) (g C m ⁻² yr ⁻¹)	1976–79 (Ref.) (g C m ⁻² yr ⁻¹)
Primary production	1570* (4, 35)	1351* (1, 35, 18)
Copepod production	100†	100†
Euphausiid production	2.5†	5†
Anchovy yield	60*†	6*†
Sardine yield	0.4*	4*
Hake yield	0.05†	0.7* (27)
Bacterioplankton production	215† (7)	22†
Microbenthic production	} 2*† (10)	9*† (9)
Meiobenthic production		2.5*† (9)
Macrobenthic production	0.5*† (11)	0.1†
Detrital carbon production	320† (4)	720†
Sinking loss	105†	698†
Sediment storage/export	82†	591†

* Measured; † calculated.

may be exported to the continental slope as phytodetrital carbon from these shelves, thus constituting a major organic sink of the global CO₂ budget³⁰.

An annual detrital carbon loss of ~500 g C m⁻² yr⁻¹ at 15°S, and a C/N ratio of 5/1 implies that the particulate nitrogen export in the form of detritus is 100 g N m⁻² yr⁻¹, in contrast to ~1 g N m⁻² yr⁻¹ in the form of fish yield (Fig. 1b). To maintain a nitrogen mass balance in an annual steady state of the food web on this Peru shelf, ~100 g N m⁻² yr⁻¹ of dissolved nitrogen must at least be returned to the shelf to replace the detrital and fish nitrogen loss. Furthermore, an annual primary production of ~1,000 g C m⁻² yr⁻¹ at 15°S requires 200 g N m⁻² yr⁻¹, 40% of which could be supplied by uptake of ammonium from nitrogen recycling³¹ and 60% by physical input of nitrate from upwelling, because river runoff is negligible²⁴. An estimate of the nitrate supply can be made with a simple budget, $\delta\text{NO}_3/\delta t = w\delta\text{NO}_3/\delta z$, using an average upwelling rate, w , of 10 m day⁻¹ and an observed vertical nitrate gradient of 5 mg-atom NO₃ m⁻³ over the 50-m depth interval of the nearshore upwelling region off Peru⁴; this daily flux sums to 5 g NO₃ m⁻³ yr⁻¹. Over a 20–30-m euphotic zone, such an input of 100–150 g NO₃ m⁻² yr⁻¹ would be sufficient to balance the 'new' nitrogen demands³² of photosynthesis, fish removal and detrital loss from the shelf at 15°S.

Similar to recent changes in North Sea herring stocks²⁶, a 90% decline of anchovy production from 60 to 6 g C m⁻² yr⁻¹ off northern Peru as a result of overfishing, however, would lead to a much larger absolute change in the detrital carbon (nitrogen) flux at 10°S (a Δ of ~500 versus ~20 g C m⁻² yr⁻¹) than a similar percentage reduction from 6.0 to 0.6 g C m⁻² yr⁻¹ at

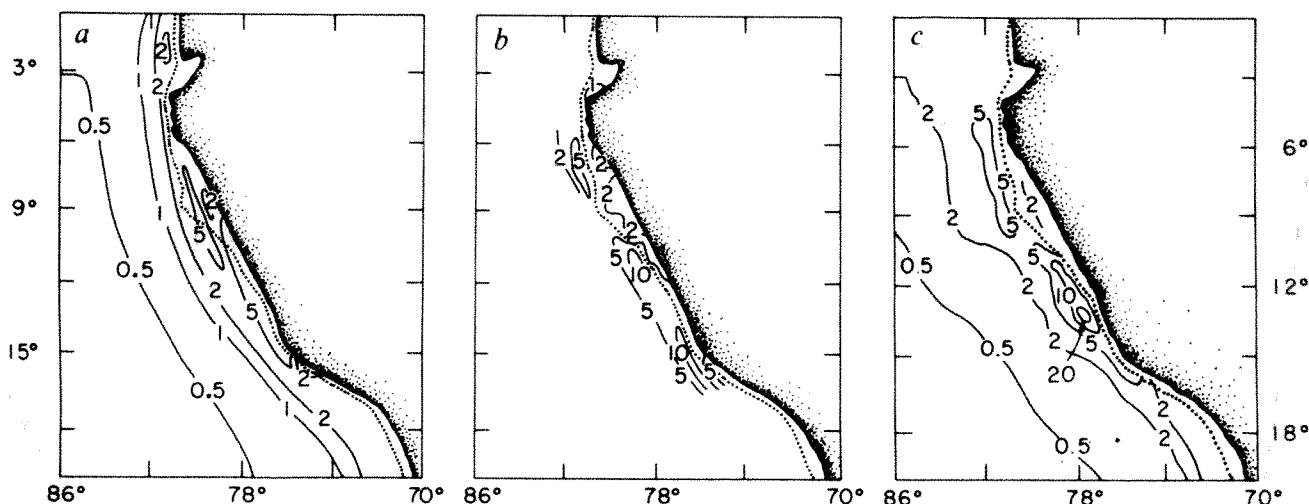


Fig. 2 The distribution of % organic carbon in the surface sediments along the Peru coast in a, 1968–69, $N = 63$; b, 1976, $N = 97$; c, 1977, $N = 56$.

15°S. The subsurface flow of water off Peru is to the south at $\sim 1\text{--}5\text{ cm s}^{-1}$ within the Peru–Chile undercurrent at the edge of the shelf. With these differences in area, food chain efficiency, and yield, the total poleward, or downstream, detrital load of the Peru system could have approximately doubled after overfishing of the anchovy stocks between 6 and 10°S. Thus part of the apparent bottom carbon transient after overfishing at 15°S, that is, a doubling between 1968 and 1978 could be the result of sinking detritus advected south within the Peru–Chile undercurrent. If the 1966–69 anchovy grazing pressure were reapplied to the shelf ecosystem at 6–10°S, presumably the upstream source of phytodetritus would decline, and the per cent sediment carbon at 15°S might also decline if the accumulated surficial bottom carbon were oxidized or further exported to the continental rise^{24,30}. Based on ²¹⁰Pb dating, cores taken at 186–370 m depth between 11 and 13°S in June 1977 suggest a sedimentation rate of 0.05–0.16 cm yr⁻¹ off northern Peru³³, whereas cores taken at 82–250 m depth near 15°S in February–March 1978 indicate a rate of 1.1–1.2 cm yr⁻¹ off Southern Peru³⁴.

Conclusions

Although the estimates in these carbon budgets are questionable it seems that most of both the larval survival and harvest of adult anchovy used to occur within a 10% area of the Peru shelf. This is about the same per cent area of the Bering Sea from which most stocks of the Alaska pollock are harvested⁵. The

hypothesis that greater changes in shelf trophodynamics have occurred within the 6–10°S area than in the 15°S region seems to reflect the impact of both overfishing by man and alteration of larval survival by nature off Peru¹. The decline of anchovy grazing pressure has apparently led to increases in plankton biomass, increased stocks of sardine and hake, increased carbon loading and sulphate reduction at the 'downstream' sea bed, and a decline in oxygen and nitrate content of the water column.

Although the anchovy fishery of the Peru upwelling ecosystem was one of the world's most 'studied', 'advised' and 'managed' coastal resource, societal conflicts led to its possible demise. Along the whole coast, the fish yield may now be only 10 tonnes km⁻² yr⁻¹, similar to that from the Mid-Atlantic Bight, North Sea and Bering Sea despite a three- to sevenfold higher primary production off Peru. Future studies will determine whether after these changes in carbon export, the carrying capacity of this system with respect to living resources will remain at present level. It is possible that the Peru upwelling ecosystem will either return to the same level of secondary production before 1972, or maintain the present decline in food chain yield such that the greater export in the form of detrital carbon rather than of fish carbon will continue.

The data collection and reduction phase of this research was supported mainly by the NSF as part of the Coastal Upwelling Ecosystem Analysis and Process and Resources of the Bering Sea programs. Additional data analysis was supported by the DOE as part of our Atlantic Coastal Ecosystem program. Drs G. Rowe and E. Suess provided sediment carbon data.

Received 30 June 1980; accepted 6 January 1981.

- Walsh, J. J. *et al. Deep-Sea Res.* **27**, 1 (1980).
- Cushing, D. H. *F.A.O. Fish Biol. Tech. Pap.* **84**, 1 (1969).
- Ryther, J. H. *Science* **166**, 72 (1969).
- Walsh, J. J. *Deep-Sea Res.* **22**, 201 (1975).
- Walsh, J. J. in *Analysis of Marine Ecosystems* (ed. Longhurst, A. R.) 159 (Academic, New York, 1980).
- De Mendiola, B. R. in *The Early Life History of Fish* (ed. Blaxter, J. H. S.) 277 (Springer, Berlin, 1974).
- Walsh, J. J. in *Estuarine Research* Vol. 1 (ed. Cronin, L. E.) 617 (Academic, New York, 1975).
- Codispoti, L. A. & Packard, T. T. *J. mar. Res.* **39**, 453 (1980).
- Rowe, G. T. in *Bioproductivity of Upwelling Ecosystems* (eds Barber, R. T. & Vinogradov, M. Y.) (Springer, Berlin, in the press).
- Pamatmat, M. *Limnol. Oceanogr.* **16**, 536 (1971).
- Rowe, G. T. *Inv. Pesq.* **35**, 127 (1971).
- Smith, K. L. *Ecology* **54**, 1065 (1978).
- Smith, K. L., Rowe, G. T. & Nichols, J. A. *Estuar. Coast. mar. Sci.* **1**, 65 (1973).
- Walsh, J. J. & Dugdale, R. C. *Inv. Pesq.* **35**, 309 (1971).
- Fiadario, M. & Strickland, J. D. H. *Deep-Sea Res.* **26**, 187 (1965).
- Goering, J. J. & Pamatmat, M. *Inv. Pesq.* **35**, 233 (1971).
- Dugdale, R. C., Goering, J. J., Barber, R. T., Smith, R. L. & Packard, T. T. *Deep-Sea Res.* **24**, 601 (1977).
- Project ICANE (L.A.E. DOE Dept. Ser. BI-R-78-6, Bedford Institute of Oceanography, Halifax, 1978).
- Shushkina, E. A., Vinogradov, M. Y., Sorokin, Y. I., Lebedeva, L. P. & Mikheyev, V. N. *Oceanology* **18**, 579 (1978).
- Muller, P. J. & Suess, E. *Deep-Sea Res.* **26**, 1347 (1979).
- Longvineko, N. V. & Romankevich, E. A. *Litol. Pol. Isk.* **1**, 3 (1973).
- Rodionova, K. F., Romankevich, E. A., Duzhikova, T. N. & Teklova, M. S. *Vnigni, TRUDY 138, Geochim. Sbornik* **5**, 118 (1973).
- Suess, E. & Muller, P. J. *Proc. Cent. natn. Rech. Scient.* (in the press).
- Walsh, J. J., Premuzic, E. T. & Whitledge, T. E. in *Ecophysiology* (ed. Nihoul, J. C. J.) (Elsevier, Amsterdam, in the press).
- Brinton, E. *Bull. Scripps Inst. Oceanogr.* **8**, 51 (1962).
- Burd, A. C. *Rapp. Reun. Cons. Int. Explor. Mer.* **172**, 137 (1978).
- Bruzhinin, A. D. & Pshenichny, B. P. in *Bioproductivity of Upwelling Ecosystems* (eds Barber, R. T. & Vinogradov, M. Y.) (Springer, Berlin, in the press).
- Dugdale, R. C. & Goering, J. J. *Anton Brunn Rep.* **4**, 53 (1970).
- Ryther, J. H., Menzel, D. W., Hulbert, E. M., Lorenzen, C. J. & Corwin, N. *Inv. Pesq.* **35**, 43 (1971).
- Walsh, J. J. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 497 (Plenum, New York, 1980).
- Whitledge, T. E. in *Upwelling Ecosystems* (eds Boje, R. & Tomczak, M.) 90 (Springer, Berlin, 1978).
- Dugdale, R. C. & Goering, J. J. *Limnol. Oceanogr.* **12**, 196 (1967).
- DeMaster, D. J. thesis, Yale Univ. (1979).
- Henrichs, S. M. thesis, MIT-WHOI (1980).
- Barber, R. T. & Smith, R. L. in *Analysis of Marine Ecosystems* (ed. Longhurst, A. R.) 31 (Academic, New York, 1980).
- Cowles, T. J., Barber, R. T. & Guillen, O. *Science* **195**, 285 (1977).

In vivo sequence requirements of the SV40 early promoter region

Christophe Benoist* & Pierre Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg, France

To investigate the sequences necessary for proper initiation of transcription of SV40 early genes, we have constructed several deletion mutants in the promoter region. The TATA box region is apparently involved in fixing initiation precisely within a narrow area, but is dispensable for gene expression, while the sequences located more than 150 base pairs upstream are indispensable.

IN eukaryotic systems, the identity and even the existence of the DNA sequences which function as promoters for RNA polymerase B are as yet uncertain. DNA sequences of eukaryotic

genes (see refs 1, 2 for reviews) and their adjacent regions nevertheless contain an A+T-rich region, typically 5'TATA_AT_A3' at a strongly conserved distance 30 base pairs from the nucleotide coding for the first base of the mature mRNA, called the cap site and presumably the point at which transcription begins³. The A+T-rich region, known as the

* Present address: Stanford University Medical School, Department of Genetics, Stanford, California 94305, USA.

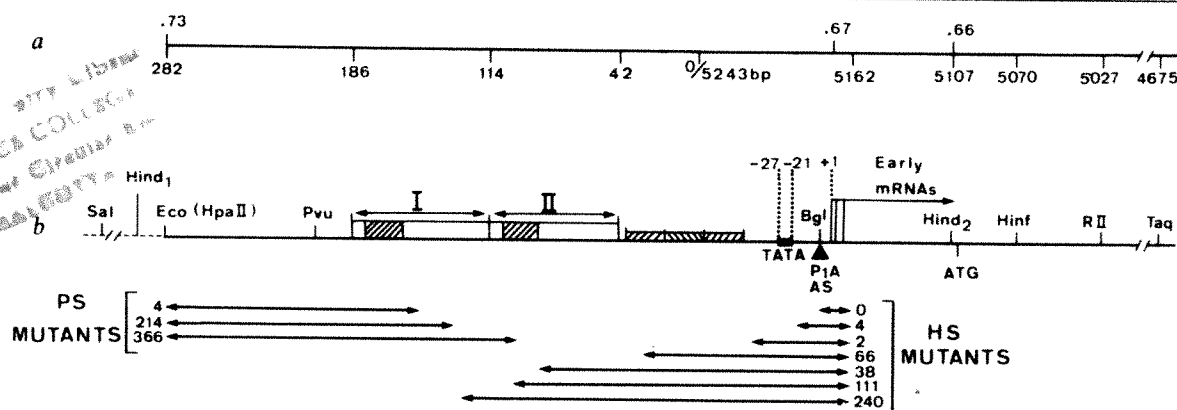


Fig. 1 a, Scale in SV40 map units (above the line) or in base pairs (bp), the latter according to the numbering system¹³ of Reddy *et al.*⁴¹ modified to take into account a previously overlooked 17 base pairs²⁹. b, Physical map of the SV40 early promoter region in the pSV1 recombinant¹³, and schematic representation of the deletions present in some of the mutants analysed in this study. SV40 sequences are represented by a solid line, pBR322 sequences by a dashed line. Sal, Hind, Eco, Pvu, Bgl, Hinf, RII and Taq represent sites for restriction enzymes *Sal*I, *Hind*III, *Eco*RI, *Pvu*II, *Bgl*I, *Hinf*I, *Eco*RII and *Taq*I, respectively (only the relevant sites are shown). *Eco*RII indicates the *Eco*RI site of pSV1 created by the blunt-end ligation of the repaired SV40 *Hpa*II site into the repaired *Eco*RI site of pBR322¹³. The position of the 5' ends of wild-type SV40 early mRNAs are shown by vertical lines, followed by a horizontal arrow which indicates the direction of transcription [+1 corresponds to the first cap site at position 5,172 (refs 19, 20)]. The position of ATG translation initiation codon is shown next to the restriction site *Hind*2. The small solid box indicates the position of the TATA box (position -21 to -27 from position +1). Large boxes represent the 72-base pair repeats I and II, and the dashed areas within correspond to the G+C-rich blocks. The three small dashed boxes are the shorter, G+C-rich repeats (see Fig. 2a). The position of the deletions in mutants P1A and AS is shown under the *Bgl*I site. Horizontal arrows define the extent of the deletions, as determined by DNA sequence analysis, in representative HS and PS mutants. For the construction of P1A, pSV1 was linearized by partial cleavage with *Bgl*I (there are three *Bgl*I sites in the pBR322 portion of pSV1 along with the unique site of interest in the SV40 moiety). The *Bgl*I 3'-protruding extremities were removed with *Escherichia coli* DNA polymerase I (Biolabs), and the plasmids circularized by blunt-end ligation with T4 DNA ligase. P1A was isolated by screening recombinants for the absence of the *Bgl*I site in the SV40 sequences. The deletion in P1A led to the generation of a *Sac*II site at this position. We took advantage of this unique site to cleave P1A by *Sac*II, and, by the same procedure as that used for construction of P1A, isolated the mutant AS. To construct the mutants of the PS series, pSV1 was linearized at the unique *Eco*RI site, and treated with *Bal*31 nuclease (BRL)^{13,42} to remove up to 200 base pairs from the extremities thus generated. The DNA was then treated with *Sal*I and ligated to an *Eco*RI-*Sal*I fragment from pBR322. Selection of tetracycline-resistant bacteria allowed the isolation of mutants bearing deletions on the SV40 side of the *Eco*RI site only. The construction scheme of HS mutants has been reported in detail previously¹³.

Goldberg-Hogness or the TATA box, has been regarded as a promoter sequence analogous to the Pribnow box sequence of prokaryotic organisms. Nevertheless, genes (such as those coding for papovavirus late function and adenovirus-2 72,000 molecular weight (72K) DNA binding proteins)⁴⁻⁸ lack the TATA box, so that its proposed function remains conjectural.

Using *in vitro* transcription systems that allow RNA polymerase B to transcribe cloned DNA templates^{9,10}, the TATA

box has been shown to be involved in promoter function *in vitro*^{2,11,12}, but *in vivo* studies suggest that these results should be extrapolated with caution. Thus for SV40 early genes, we and others¹³⁻¹⁵ have shown that removal of all sequences upstream from the cap site, but not the mere removal of the TATA box¹³, abolishes expression. Similarly, the TATA box is not essential for expression of the sea urchin histone H2A gene injected into *Xenopus* oocytes although its absence leads to the use of several new cap sites¹⁶.

Fig. 2 Sequence of the SV40 deletion mutants. The sequence corresponds to the DNA non-coding strand of the early promoter region which was cloned in pSV1, and is identical to that of wild-type SV40 strain 776^{29,41,43}, except for an 8-base pair deletion (positions 116-108, see text). The numbering system is as in Fig. 1. The 72-base pair repeats are shown in two dashed boxes, the TATA box in a solid box. Horizontal lines above the sequence delimit the five G+C-rich blocks (see text). a, Sequence of mutants in the HS group, which all share a common end point at position 5,162 indicated by a vertical line and a large arrow. The other deletion end points are shown by thinner arrows (and dotted lines for those which have been sequenced). Deletions present in mutants P1A and AS around the *Bgl*I site are indicated by double pointed arrows. b, Sequence of the mutants in the PS series. All deletions originate at the *Eco*RI site (Fig. 1) and extend to the point defined by thin arrows. The extent of the deletions was established for all mutants by DNA sequencing except for mutants HS102, HS40, PS2 and PS359, where the size (± 8 base pairs) of the deletions were estimated by gel electrophoresis after *Hind*III digestion. Sequences were obtained by the technique of Maxam and Gilbert²¹ after 5'-end labelling at the *Hind*III site 2 (Fig. 1) and secondary cleavage with *Eco*RI (P1A, AS and HS mutants) or by end labelling at the *Hind*III site 1 and secondary cleavage with *Bgl*I (PS mutants). The exact sequence of some of the mutants is shown in Fig. 6a.

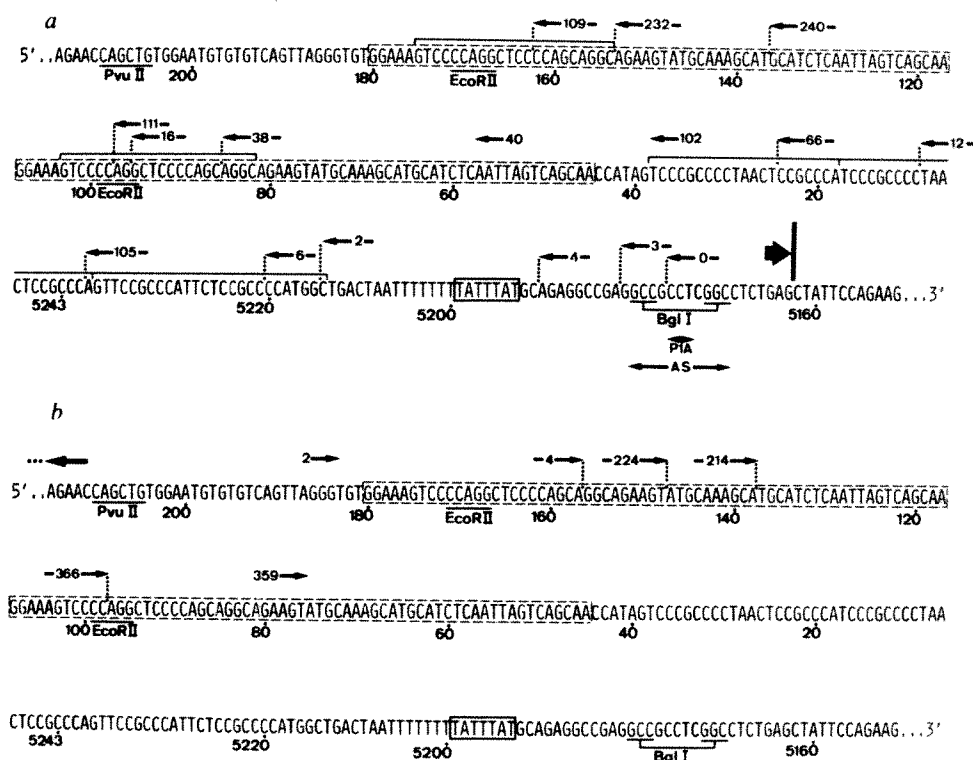
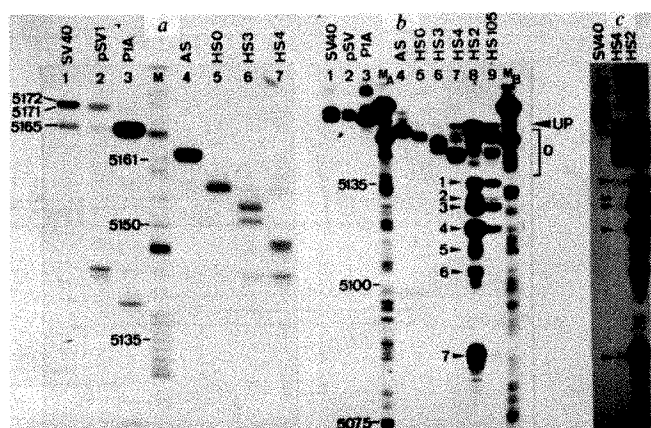


Fig. 3 Mapping of the 5' ends of SV40 early mRNAs present in several rat (3T3) fibroblast transformed cell lines (see text). Cytoplasmic RNA was extracted from such cloned cell lines by NP40 lysis and phenol-chloroform extraction, and poly(A)⁺ RNA selected by oligo(dT)-cellulose chromatography. *a*, Mapping by reverse transcriptase primer extension. A *Hinf*I-*Hind*III primer (positions 5,070–5,107, Fig. 1) ³²P-5'-end labelled at its *Hinf*I extremity, was purified by electrophoresis on an 8% acrylamide/7 M urea gel. It was then denatured in H₂O for 2 min at 100 °C, quick-frozen in liquid nitrogen and thawed in 50 mM Tris-HCl pH 8, 10 mM MgCl₂ and 140 mM KCl in the presence of poly(A)⁺ RNA (2–5 µg). After 5 min at 65 °C dATP, dGTP, dCTP and TTP (0.5 mM each) and β-mercaptoethanol (20 mM) were added and the samples were incubated for 2 h at 37 °C in the presence of 4 units of reverse transcriptase. Samples were diluted to 0.1 M NaOH, 3 mM EDTA and 5 M urea, and aliquots electrophoresed on 8% acrylamide/7 M urea thin gels⁴⁴. RNAs were from cell lines transformed with wild-type SV40 or recombinant mutants as indicated. 'M' is a sequencing ladder obtained by Maxam and Gilbert²¹ cleavage (A + C) of a *Hinf*I-*Eco*RI fragment of pSV1 (positions 5,070–282) end labelled at its *Hinf*I extremity. The positions of the '5'-end bands' are given in Fig. 6 legend. Aside from the major cap sites, a few other bands can be seen (see lanes 2 and 3, for example) that probably correspond to reverse transcriptase 'false stops' as they are not obtained by S₁ mapping. *b*, S₁ nuclease mapping of the 5' ends of the RNAs (as indicated). The coding strand of SV40 fragment *Eco*RII G (positions 5,027–99) was 5'-end labelled and purified by polyacrylamide gel electrophoresis as described in ref. 21. After denaturation for 5 min at 85 °C, the probe was hybridized to RNAs (1–3 µg) from transformed cell lines for 12 h at 42 °C in 50% formamide, 10 mM PIPES pH 6.4 and 0.4 M NaCl. Digestion with S₁ nuclease (Miles) and subsequent steps were as in ref. 45. S₁-resistant hybrids were analysed by electrophoresis on an 8% acrylamide/7 M urea thin gel. M_A and M_B: sequencing ladder from an *Eco*RII/*Bgl*I (5,027–5,177) cleaved by the technique of Maxam and Gilbert [M_A, A + C; M_B, C + T]. Arrows 1–7 indicate the common bands found with RNAs from all lines transformed with HS mutants lacking the TATA box (positions 5,134, 5,127, 5,124, 5,117, 5,112, 5,103 and 5,085). The 'O' bracket shows the area where initiation varies from one mutant to the next (see text). The 'UP' band corresponds to RNA species initiating upstream from the position of the deletion (see text). *c*, Same as *b*, where a longer exposure reveals the presence of bands 1–4 and 7 obtained with HS4 RNA (arrowheads), but not with SV40 RNA. Numbering of lanes as in *b*.



The SV40 early gene products (small t and large T antigens), which initiate viral DNA replication in permissive monkey cells, stimulate various host cell functions and transform non-permissive systems (see ref. 17 for review), are coded for by different mRNAs derived from a common primary transcript by different splicing patterns, and have identical 5' and 3' ends^{18,19}. The major 5' ends of SV40 early mRNAs map at a few closely positioned cap sites, around positions 5,165 and 5,171 on the SV40 sequence (refs 19, 20; see also Figs 1, 6). Upstream from the cap sites are located, successively, a 27-base pair palindrome, a TATA box (TATTTAT, at position -21 to -27 base pairs from the most upstream of the cap sites), a nearly perfect triple tandem of a highly G + C-rich 21-base pair sequence, and a perfect tandem repeat of a 72-base pair sequence (-116 to -188 and -189 to -261 base pairs from the cap site) (see Figs 1, 2).

We have previously reported¹³ that T-antigen synthesis can be induced by mutant plasmids containing an SV40 early region from which the TATA box has been deleted. We have now analysed the 5' ends of RNA species synthesized on such mutant templates, and conclude that initiation at the wild-type cap sites requires the presence of the TATA box. In the absence of the TATA box, transcription begins at several new sites. Moreover, we find that the expression of the SV40 early genes is severely impaired by deletions in the region of the 72-base pair repeats, which accordingly seems to have a crucial and independent role in initiation of transcription.

Deletion mutants of SV40

The mutants described here were derived from the plasmid pSV1¹³, which contains the entire early gene segment (as the large *Bam*HI-*Hpa*II fragment of SV40), cloned into the *Eco*RI site of pBR322. From this, we have constructed three groups of deletion mutants (see Fig. 1 legend). After preliminary screening by restriction enzyme digestion, the extent of the deletions was established (in most cases, see Fig. 2) by DNA sequencing²¹.

The first group of mutants is represented by plasmids P1A and AS, with deletions between the TATA box and the cap sites of 3 and 9 base pairs, respectively (Figs 1, 2a). Figure 1 depicts schematically the position of the deletions in the second (PS) and third (HS) groups of mutants. Deletions in the PS series share an end point at the *Eco*RI site at the boundary of the SV40 and pBR322 sequences in pSV1, and extend progressively further downstream towards the early genes. The same pBR322 sequences are always present at the boundary, which greatly facilitates screening and eliminates one source of variability in

studying the effect of deletions in the SV40 sequence. Deletions in the HS mutants all extend upstream from position 5,162 (ref. 13), and thus delete the cap site and usually the TATA box, which is nevertheless left intact in HS0, HS3 and HS4 (Figs 2, 6). On mutant sequences, all have an 8-base pair deletion at the junction between the two 72-base pair repeated sequences (positions 116–108), apparently because the SV40 fragment cloned in pSV1 is derived from a variant of SV40 strain 776 (Fig. 2).

In vivo expression of SV40 early genes in mutant plasmids

Our demonstration that deletion of all sequences upstream from the cap site abolishes SV40 early gene expression¹³, shows that some sequence element between the TATA box and *Hpa*II site

Table 1 *In vivo* expression of the deletion mutants

Short-term		Long-term	
Plasmids	T antigen-positive cells per 2 × 10 ⁴ cells	Plasmids	Foci per plate
pSV1	105	pSV1	132
AS	72	HS4	115
HS4	91	HS2	118
HS2	78	HS105	72
HS12	17	HS102*	17
HS66	15	HS38*	21
HS40	6	HS111*	7
HS38	8	HS240*	1
HS111	0	PS214*	15
PS224	12	PS366*	13
PS366	0		

Short-term experiments: Plasmid DNA (0.2 µg) was diluted in 200 µl of Dulbecco's modified Eagle's medium (DMEM), followed by the addition of 200 µl of a 1 mg ml⁻¹ solution of DEAE-dextran (Sigma) in DMEM. Sub-confluent CV1 monkey cells, grown in 60-mm Petri dishes¹³, were washed twice with phosphate-buffered saline (PBS); the DNA solution was added and, after incubation for 15 min at room temperature, the cells were washed once with PBS and once with complete medium. After addition of medium, they were incubated at 37 °C. Indirect immunofluorescence was performed 40 h after transfection, essentially as described in ref. 39, 10 µl of the antibodies¹³ being spotted directly on to the cell monolayer fixed to the dish with acetone/methanol (3:7 v/v). T antigen-positive cells were counted under a fluorescence microscope. Long-term experiments: Transformation of non-permissive cells. Fisher rat fibroblasts (FR3T3), grown in 100-mm Petri dishes¹³, were transfected with 1 µg of plasmid DNA, as described in ref. 40. Dense foci overgrowing the cell monolayer were counted 3 weeks after transfection. Asterisks indicate that foci were counted 4 weeks after transfection and that they were checked by immunofluorescence to eliminate foci due to spontaneous transformation (5–10 per dish). All the data shown in this table represent results obtained in typical experiments. The absolute figures are only indicative, however, as they may vary by as much as 50% between experiments.

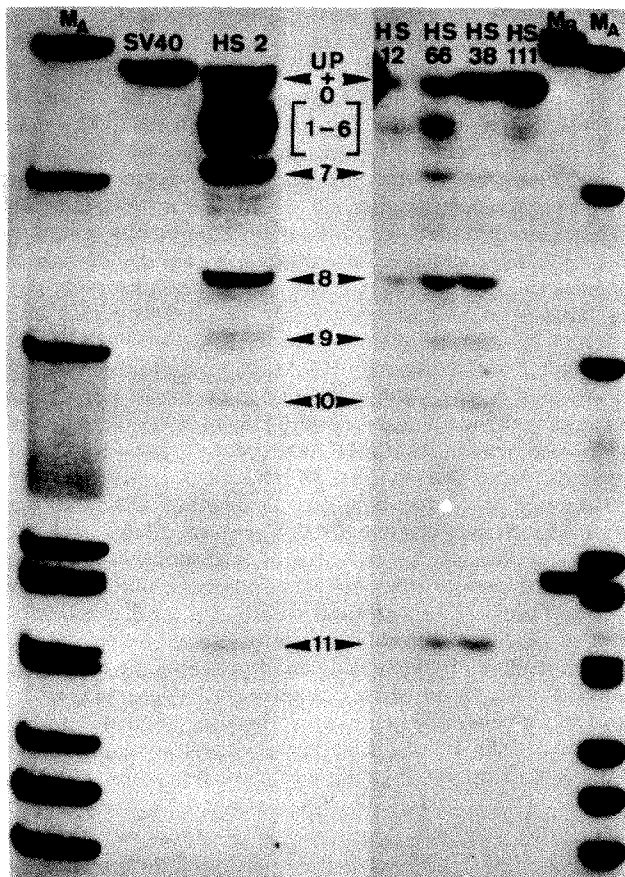


Fig. 4 S_1 nuclease mapping showing the full extent of the 5' ends of RNAs from cells transformed with HS mutants lacking the TATA box. The experimental protocol is the same as that described in Fig. 3b legend (4 μ g of poly(A)⁺ RNA for SV40 and HS2; HS12, 8 μ g; HS66, 38 μ g; HS111, 15 μ g); the probe used was the coding strand of the *TaqI*-*BglI* fragment of SV40 (positions 4,675–5,180), 5'-labelled at its *TaqI* extremity. S_1 -resistant hybrids were analysed on a 5% acrylamide/7 M urea thin gel. M_A : 5'-end labelled fragments from a *HpaII* digest of pBR322 (sizes from top in base pairs: 527, 404, 309, 242, 238, 217, 201, 190, 180). M_B : 5'-end labelled fragments from a *HinfI* digest of SV40. The origin of the RNAs is indicated at the top of the lanes. The resolution on a 5% polyacrylamide gel is not sufficient to separate the band mapping at the position of the deletion (UP) from those immediately below (region O) (see Fig. 3b). Similarly, bands 1–6 are poorly separated. Bands 7–11 correspond to 5' ends of RNA species mapping at positions (± 10 base pairs) 5,085, 5,023, 4,985, 4,962 and 4,897, respectively. (Only the major bands 1–11 are numbered, but a large number of more minor cap sites are also present, for example, HS2 RNA, where at least 10 such bands can be seen between bands 7 and 11 on the original autoradiogram.)

(position 282, Fig. 1) is essential for transcription. We have therefore analysed the *in vivo* expression of the early genes of HS and PS mutants (Fig. 2).

Two approaches were followed (Table 1). In short-term experiments, gene expression was detected by T antigen-specific immunofluorescence 40 h after transfection of permissive CV1 monkey cells with the recombinant plasmids. In long-term experiments, we assayed the ability of the deletion mutants to induce morphological transformation of rat 3T3 fibroblast cells (tested by the appearance of T antigen-positive dense foci overgrowing the cell monolayer).

In the short-term experiments, HS mutants (Fig. 2a) up to and including HS38 induced a T antigen-positive immunofluorescence pattern, with an intensity unaffected by deletion of the TATA box (HS2 and 6), but inversely proportional to the length of the deletions (HS66, 40 and 38). On the other hand, the responses with HS16, 111, 240, 232 and 109 were not unambiguously positive. With the sequence of PS mutants (Fig. 2b), gene expression was allowed down to those in PS224 and 214, but with these two mutants the fluorescence intensity was lower than that obtained with the wild-type recombinant pSV1; in contrast, PS366 and 359 gave negative results. We have also assayed by blot hybridization²² for SV40-specific early RNAs in

poly(A)⁺ RNA extracted from CV1 cells 40 h after transfection with various plasmids, and have found none in cells when the mutants were negative by immunofluorescence (not shown). The sensitivity of such experiments is low, however, and it is possible that a low level of transcription (<10% of that obtained with pSV1) occurs with mutants which do not induce T-antigen immunofluorescence.

These experiments therefore show that the region of the 72-base pair repeats (Figs 1, 2) is necessary for efficient expression of the early genes. However, morphological transformation studies were somewhat more ambiguous. As expected, rat 3T3 cells were transformed by HS mutants up to HS38, but the efficiency of transcription was decreased with mutants such as HS40 or HS38. Using RNA blot analysis, we have quantitated the amount of early RNAs present in lines of such transformants. Like Gluzman *et al.*¹⁴, we found that the levels of these RNAs were higher in cells transformed with mutants bearing deletions downstream from the TATA box than in cells transformed with pSV1 (although the increase varied from 2- to 10-fold with different lines; our unpublished results). These levels were not affected by further deletion of the TATA box, but decreased with HS mutants bearing longer deletions (not shown). However, in contrast to the short-term experiments, T antigen-positive transformed foci were induced by HS111, HS240 and PS366, although in reduced numbers (5–10 foci per μ g of DNA per 10^6 cells compared with 50–100 for pSV1 and short deletion HS mutants) and more slowly (25–30 days compared with 10–14 days).

Transcription initiation on mutants retaining the TATA box

Using both S_1 nuclease mapping with single-stranded ³²P-5'-end-labelled probes²³ and reverse transcriptase primer extension mapping¹⁹ (see Figs 1, 3), we have investigated the 5' ends of SV40-specific RNAs present in cloned rat fibroblast cell lines transformed by HS mutants. We believe both types of experiments to be necessary because of the risk of artefacts, particularly with the primer extension method (see Fig. 3 legend). We have assigned (Fig. 6a) the major 5' ends of SV40 early mRNAs to positions 5,172, 5,171 and 5,165 (both in RNA isolated 32 h after infection by wild-type virus²⁴ and in lines transformed by wild-type SV40 or pSV1; Fig. 3a and unpublished results). However, because results obtained by the two methods often differ by 1 or 2 base pairs (compare, for instance, Fig. 3a with Fig. 5a of the accompanying paper²⁴), and because of discrepancies in previous reports (refs 18–20 and Y. Groner, personal communication), we do not consider these assignments definitive.

Figure 3a and b give the 5'-end analysis of RNAs transcribed from deletion mutants retaining the TATA box. The 3-base pair deletion in P1A results in initiation mainly at position 5,165, with almost complete disappearance of the two wild-type 5' ends upstream at positions 5,172 and 5,171. On mutants with longer deletions downstream from the TATA box (AS, HS0, HS3, HS4), transcription begins further down, with the 5' ends always mapping at about the same distance from the TATA box (27–34 base pairs from the first T, Fig. 6a). This shift seems to be independent of the presence (P1A, AS) or absence (HS0, HS3, HS4) of the wild-type cap sites. In those mutants which do conserve the TATA box, there are very few cap sites.

Transcription initiation on mutants lacking the TATA box

The clearest consequence of the removal of the TATA box is the generation of multiple cap sites, as shown for mutants HS2 and HS105 (Fig. 3b). Identical bands (arrows 1–7 in Fig. 3b) were observed with all the mutants lacking the TATA box, but with varying intensities (see below). Between band 1 and that corresponding to the end point common to all deletions (arrowhead 'UP'), patterns varied between mutants (region O in Fig. 3b) although some cap sites were common to several mutants (Fig.

5). We investigated this multiplicity further by S_1 mapping using a single-stranded probe end labelled at the *TaqI* site (position 4,675, Fig. 1), and found that in the absence of the TATA box, initiation occurs up to 270 base pairs downstream from the normal 5' ends (Fig. 4, band 11). These cap sites are not usually seen with wild-type SV40 RNA (Fig. 4), but it is just possible to detect them using very large amounts of such RNA (not shown). As the probes used in the experiments shown in Figs 3b and 4 were derived from wild-type SV40 DNA, RNA species from HS mutants and with their 5' ends mapping upstream from the position of the deletion lead to the appearance of a band ('UP' in Figs 3b, 4) at that position. With mutants such as HS2 and 105 these RNAs are relative minor, with most of the initiation occurring downstream.

Figure 4 shows the effect of increasingly long HS deletions including the TATA box and sequences upstream. Although identical bands are found with mutants HS2, 12, 66 and 38, the ratio of the intensity of these bands varies, the lower bands becoming more prominent with increasing deletion. For example, whereas bands 1–6 are most important in HS2 RNA, they are barely detectable in HS38 RNA, while the relative intensities of the lower bands 8–11 increase (see Fig. 6b for quantitation). It thus appears (similar results were obtained reproducibly in several experiments) that, although identical cap sites are shared by various HS mutants, their use is influenced by their distance from sequences found upstream from the HS38 deletion end point.

HS111 RNA shows a very different pattern, with the disappearance of the lower bands 8–11 and a relative reappearance of the higher bands 1–7; most of the RNA initiates upstream from the position of the deletion ('UP' in Fig. 4). This 'UP' band is also intense in HS38 RNA. We investigated the cap sites located upstream from the position of the deletion by mapping the 5'

ends of RNAs from cell lines transformed with HS2, 105, 66 (not shown), 38 and 111 (Fig. 5) with probes derived from corresponding mutant DNAs. Autoradiograms from experiments such as those shown in Figs 3, 4 and 5, were scanned and the results are summarized in Fig. 6b. With HS2, 105 and 66 RNAs the 'upstream' RNA species are heterogeneous and quite minor, their 5' ends mapping immediately upstream from the position of the deletion. For HS111, however, they are of much higher relative abundance, and are represented by a set of bands (Fig. 5) which correspond to initiation within the 72-base pair repeat 1 (Figs 1, 2). This set of bands is also observed, but at much lower relative abundance (Fig. 6b) and with a somewhat different distribution of intensities (Fig. 5), with HS38 RNA, where most of the upstream RNAs map far into the late region (X on Figs 5, 6b). Although very faint bands, perhaps corresponding to the set observed with HS111 RNA, were detected in HS2 RNA, no other HS RNAs showed any species equivalent to the 'X' bands seen with HS38 RNA.

Although we have assumed throughout that the bands obtained in S_1 mapping experiments correspond to 5' ends of the RNAs it is possible that these bands correspond to the position of acceptor sites in RNA splicing. However, multiple splicing events would then appear suddenly after the small deletion from HS4 to HS2, and the relative use of acceptor sites would be further modified by increasing deletions. Splicing can be ruled out for RNAs transcribed on mutants which conserve the TATA box and for bands 1–4 in the cases where this sequence is absent, as these 5' ends are also found by reverse transcriptase primer extension mapping (Fig. 3a and data not shown). Another interpretation is that as the RNAs studied are extracted from transformed cell lines, some of the bands seen in the S_1 nuclease mapping experiments could have been due to sequence rearrangements or integration behind cellular promoters during transformation. As such events are unlikely to occur similarly in a number of cell lines transformed by different mutants, this objection is probably not valid for bands 1–11 which are found with all mutants, nor for the set of bands found in both HS111 and HS38 RNAs, but may explain the bands found only with HS38 RNA, in region X (Fig. 5).

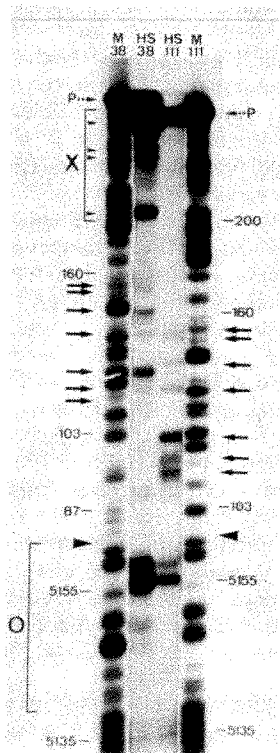
Function of the TATA box region

Our studies on the above mutants, in agreement with recent reports (refs 14, 16 and Y. Gluzman, personal communication), clearly show that the TATA box has a key role in fixing the start sites within a narrow area (we assume that the RNA 5' ends correspond to transcription start sites). Deletions in such mutants shift these sites so that initiation always occurs 27–34 base pairs downstream from the first T of the TATTTAT sequence (the slight variations in this distance and in the number of start sites presumably reflect the influence of the actual DNA sequence around the cap site^{2,24}).

When the TATA box is absent, new cap sites appear, but no TATA-like sequences are found upstream from the cap sites (except possibly for band 10, Fig. 4). We hypothesize that sequence elements, not obviously related to the canonical TATA box sequence, can functionally substitute for the TATA box to dictate the position of RNA 5' ends and might also direct initiation 30 base pairs downstream, as the first 5' end common to all TATA-minus mutants (band 1 in Figs 3b, 6) maps 30 base pairs downstream from the position of the deletion. In the wild-type situation, these substitute elements would be masked by the TATA box, which might be recognized more readily by RNA polymerase and/or other factors involved in transcription initiation. A functionally similar element might be operative in polyoma late genes and the gene for the adenovirus 72K DNA binding protein, which exhibit a microheterogeneity of mRNA 5' ends over a narrow area, but do not have TATA-like sequences at the expected position^{4,6}.

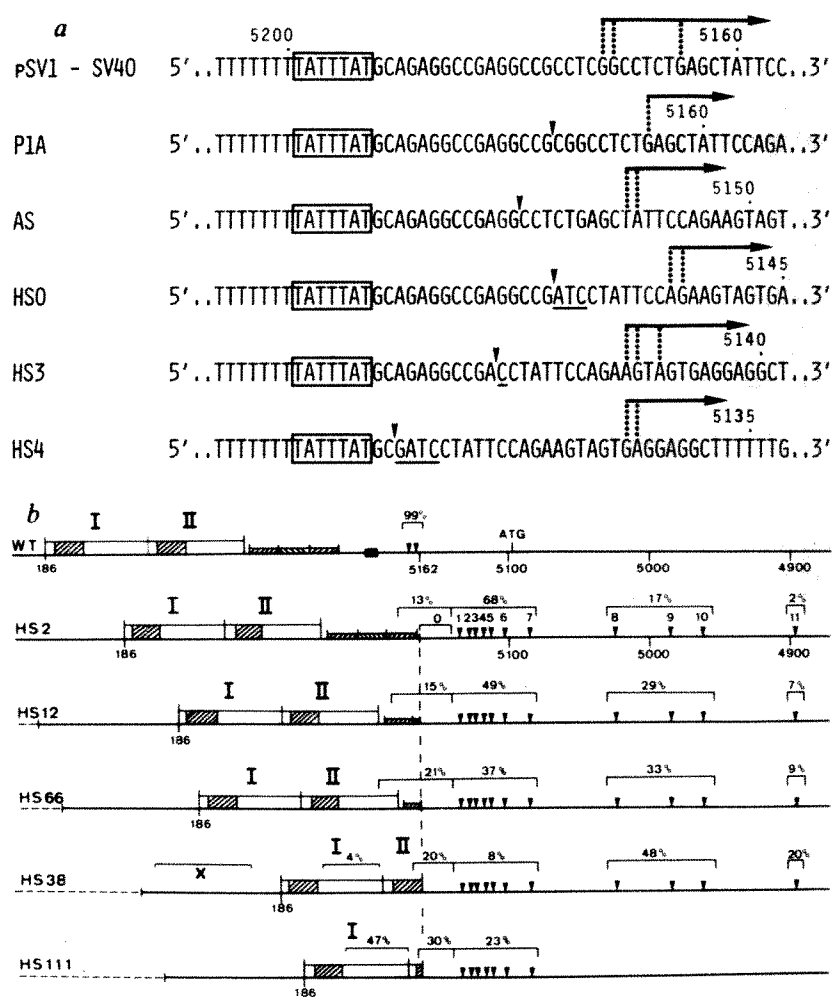
The key sequences performing the 'fixing' function for initiation at the major early cap sites must be present in HS4, but absent from HS2 (Figs 2, 3). Although the TATA box is included in this region, other sequences could influence its

Fig. 5 S_1 nuclease mapping of 5' ends of RNA species that initiate upstream from the position of the deletion, in RNA extracted from cell lines transformed with HS38 and HS111. The probes used were the *HinfI*–*EcoRI* fragments (5,070–282) derived from HS38 and HS111 DNAs, and 5'-end labelled at their *HinfI* extremity. The specific activities of the probes were different, as were the amounts of RNA used (HS38, 2 μ g; HS111, 3.5 μ g). M_{38} and H_{111} are sequence ladders obtained by (A + C) cleavage of the probes. HS38 and HS111: the probes were hybridized to RNA from corresponding cell lines, digested with S_1 nuclease and electrophoresed as described in Fig. 3 legend. The markers and experimental tracks are from different exposures of the same gel. Positions are those on the wild-type SV40 sequence (numbering as in Figs 1 and 2) and arrowheads indicate the position of the deletion (5,162). 'O' brackets the region 0 (defined in Fig. 3 legend) between the position of the deletion and band 1 at position 5,135. Arrows on the right indicate the bands obtained with HS111 RNA (positions 117, 119, 120, 135, 144, 151 and 153). These bands are observed at identical positions with HS38 RNA (arrows



on the left of the figure; the bands are faint but were clearly visible on the original autoradiogram). Bands obtained with HS38 RNA only are bracketed and labelled X. In both HS38 and HS111, a band can be seen which corresponds to full-length unreacted probe (dotted arrow 'P'). That this band is not due (at least for HS111) to RNA species initiating further upstream was established by an experiment in which the *HinfI* end-labelled probe extended for 600 base pairs into the pBR322 sequences (Fig. 1). No additional bands were observed in this experiment (not shown).

Fig. 6 Location of the 5' ends of early RNAs present in lines transformed by deletion mutants of pSV1. *a*, Position of cap sites on mutants that retain the TATA box. The sequences of the mutants are shown, aligned on the 'TATTTAT' sequence (boxed). Arrowheads indicate the position of the deletion; the few extra nucleotides which originate from the construction of HS mutants are underlined¹³. Symbols for cap sites are as in Fig. 1. Their position was derived from the results shown in Fig. 3a and other unpublished S₁ and reverse transcriptase mapping experiments and map at the following position: P1A, 5,165; AS, 5,160 and 5,161; HS0, 5,154 and 5,155; HS3, 5,149, 5,151 and 5,152; HS4, 5,145 and 5,146. As discussed in the text these positions are only tentative (± 2 base pairs). Only the major cap sites are represented, yet, as can be seen in Fig. 3a, minor cap sites exist; for example, the HS0 5' ends are also present in AS, HS3 and HS4 RNAs. *b*, Position and relative contribution of the RNA 5' ends found with HS mutants lacking the TATA box. Symbols are the same as in Fig. 1, as is the numbering system. The common deletion end point is shown by a vertical dashed line. Cap sites are shown by arrowheads and are numbered (on the second line only) as defined in Figs 3b and 4. The cap sites that vary from one mutant to the next ('region O', and the region immediately upstream from the deletion end point) are not shown individually, nor the closely spaced cap sites of mutants HS38 and HS111 that map within the 72-base pair repeat I. The brackets and the percentages above the cap sites indicate the proportion of RNA species initiating in that region, as deduced from scans of autoradiograms from several S₁ mapping experiments with wild-type probes end labelled at the *Hinf*I, *Eco*RII, *Taq*I sites (Fig. 1) or experiments with mutant probes as shown in Fig. 5. Percentages for HS38 are calculated omitting the bands in region X, which represent 45% of HS38 RNA 5' ends.



function. Indeed, bands 1-4 and 7, characteristic of mutants lacking the TATA box, can be detected faintly in HS4 RNA (Fig. 3c), but are absent in RNA transcribed on the wild-type gene or on HS mutants with shorter deletions. It is as if the TATA box in HS4 were beginning to relinquish its dominance, an interpretation supported by the relative inefficiency of HS4 DNA (see accompanying article²⁴) in directing accurate initiation of transcription *in vitro* to a site about 30 base pairs downstream.

The deletion of the TATA box does not seem to affect early gene expression quantitatively (Table 1). The possibility that down mutations would be masked if they eliminated T-antigen binding and repression²⁵⁻²⁸ probably does not apply as the T-antigen binding site II is already absent from HS4^{26,28}.

Far upstream sequences are required for transcription initiation

The most important element for SV40 early gene expression is clearly not the TATA box, for mutant PS366, which has a deletion far upstream from the TATA box, transforms rat cells very inefficiently and does not induce detectable T antigen in short-term experiments, while the expression of various mutants in monkey cells indicates that, at least in this system, the essential elements lie in the region of the 72-base pair repeats. Results obtained with mutants HS38 and PS214 suggest that deletion of one of these two repeats still allows transcription, in keeping with studies of mutants or variants of SV40 which possess only one copy of this repeat (see ref. 29 for refs). Although the lower efficiency of mutants HS66 or HS38 (Table 1) might suggest the involvement of other sequences, such as the G+C-rich 21-base pair repeats, this could also be accounted for by the position of the 5' ends of RNAs synthesized on such templates, downstream from the normal T-antigen initiation codon. The results obtained with the PS214 and PS224 are in

contrast with the aforementioned mutants and variants of SV40, whose early gene expression appears normal. This might indicate the involvement of sequences even further upstream than the 72-base pair repeats, or that pBR322 sequences have some inhibitory effect when brought closer to the gene.

Increasingly long deletions (HS2, HS12, HS66, HS38) affect the relative intensities of the common bands 1-11 with a pattern (Fig. 6b), suggesting that initiation takes place predominantly at a fixed distance (broadly speaking 150-200 base pairs) from sequences located upstream from the deletion endpoint of HS38. This pattern is drastically altered with mutant HS111, which bears a deletion only 12 base pairs longer than that of HS38. In addition, HS111 does not induce detectable T-antigen synthesis in CV1 cells, whereas HS38 does. There are five sets of highly G+C-rich 20-base pair sequences in the early SV40 promoter region (Figs 1, 2). The G+C-rich sequence of the 72-base pair repeat II is almost eliminated in HS111, but essentially left intact in HS38. We suggest that this sequence is involved in controlling the pattern of initiation in mutants HS2 to HS38; HS111 transcription would then be under the control of the same G+C-rich sequence of repeat I. If this is the case, the effect of the G+C-rich sequences must be complex: predominant initiation occurs in HS111 only 10-60 base pairs downstream from the G+C-rich sequence of repeat I, whereas it occurs ~200 base pairs downstream from its counterpart in repeat II, when this is present (HS2 to HS38). The relatively constant amount of RNA species initiating in region 0 or immediately upstream (Fig. 6b) is also puzzling.

We have observed that mutants which do not support early gene expression in CV1 cells, such as HS111, HS240 and PS366, can, albeit inefficiently, nevertheless transform rat fibroblast cells. As these transformants could reflect a low level of transcription, not detected in short-term experiments because of a lower sensitivity, we are now investigating the extent of sequences necessary for such transformation (bearing in mind

the possibility that template rearrangements could be involved; see above).

We do not know how sequences lying more than 150 base pairs from the cap sites exert their influence. The structure of SV40 chromatin is known to be in a very 'open' conformation in the early promoter region³⁰⁻³⁴. Moreover, the chromatin structure around the 5' end of several 'active' cellular genes is extremely sensitive to DNase I (refs 35, 36 and M. Bellard, personal communication). It is possible that the upstream sequences somehow induce this peculiar chromatin structure, which in turn would allow the RNA polymerase to interact, directly or indirectly, with the TATA box. Alternatively (or concomitantly), these sequences could be part of the RNA polymerase binding site²⁴, or interact with proteins exerting a positive control over transcription.

Can the organization found for the SV40 early promoter region be generalized to other systems? There is no clear homology between the SV40 sequence upstream from the TATA box and regions upstream from the 5' end of other genes. However, both polyoma early genes (R. Kamen, personal communication) and the sea urchin H2A gene (R. Grosschedl and M. Birnstiel, personal communication) have sequences lying upstream from the TATA box, in a position roughly analogous to that of the 72-base pair repeats in SV40, which are essential

for gene expression. Thus, it is possible that sequence elements will be found in other genes that have an analogous role to that of the 72-base pair repeat region in SV40.

In conclusion, if a eukaryotic promoter region is defined as being composed of all the sequences necessary for proper initiation of transcription *in vivo*, the SV40 early promoter has (at least) three components: the TATA box and cap site regions which accurately position initiation of transcription, and the sequences in the region of the 72-base pair repeats which allow, within the area they potentiate, initiation of transcription and TATA box function. The requirement for the far upstream sequences, which have no counterpart in prokaryotes, may reflect the importance of chromatin structure in the control of gene expression in eukaryotes^{37,38}.

We thank Drs J. Beard and R. Wilsnack for gifts of reverse transcriptase and antiserum, Drs R. Breathnach, K. O'Hare and D. Mathis for helpful advice, K. Dott, J. M. Garnier, M. C. Gesnel, L. Thia-Toong and Dominique Schmitt for technical assistance, and B. Boulay for photography. This work was supported by grants from the INSERM (ATP 72.79.104, CL 801029 and CL 801033), the CNRS (ATP 'Biologie Moléculaire du Gène'), the Fondation Française pour la Recherche Médicale and from the Fondation S. and C. Del Duca.

Received 1 October 1980; accepted 22 January 1981.

- Breathnach, R. & Chambon, P. A. *Rev. Biochem.* (in the press).
- Corden, J. *et al. Science* **209**, 1406-1414 (1980).
- Ziff, E. & Evans, R. *Cell* **15**, 1463-1475 (1978).
- Flavell, A., Cowie, A., Arrand, J. & Kamen, R. *J. Virol.* **33**, 902-908 (1980).
- Haegeman, G. & Fliers, W. *Nucleic Acids Res.* **5**, 2359-2371 (1978).
- Baker, C., Herisse, J., Courtois, G., Galibert, F. & Ziff, E. *Cell* **18**, 569-580 (1979).
- Groner, Y., Carmi, P. & Aloni, Y. *Nucleic Acids Res.* **4**, 3959-3968 (1978).
- Ghosh, P., Reddy, V., Swinscoe, J., Lebowitz, P. & Weissman, S. *J. molec. Biol.* **126**, 813-846 (1979).
- Weil, P., Luse, D., Segall, J. & Roeder, R. *Cell* **18**, 469-484 (1979).
- Manley, J., Fire, A., Cano, A., Sharp, P. & Gelfand, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3855-3859 (1980).
- Wasylyk, B., Kedinger, C., Corden, J., Brison, O. & Chambon, P. *Nature* **285**, 367-373 (1980).
- Wasylyk, B. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Benoist, C. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3865-3869 (1980).
- Gluzman, Y., Sambrook, J. & Frisque, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3898-3902 (1980).
- Mueller, C., Graessmann, A. & Graessmann, M. *Cell* **15**, 579-586 (1978).
- Grosschedl, R. & Birnstiel, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1432-1436 (1980).
- Acheson, N. in *Molecular Biology of Tumor Viruses* Pt 2 (ed. Tooze, J.) 125-204 (Cold Spring Harbor, New York, 1980).
- Thompson, J., Radonovich, M. & Salzman, N. *J. Virol.* **31**, 437-446 (1979).
- Reddy, V., Ghosh, P., Lebowitz, P., Piatak, M. & Weissman, S. *J. Virol.* **30**, 279-296 (1979).
- Haegeman, G. & Fiers, W. *J. Virol.* **35**, 955-961 (1980).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-580 (1980).
- Alwine, J., Kemp, D. & Stark, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
- Weaver, R. & Weissmann, C. *Nucleic Acids Res.* **7**, 1175-1193 (1979).
- Mathis, D. J. & Chambon, P. *Nature* **290**, 310-315 (1981).
- Reed, S., Stark, G. & Alwine, J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3083-3087 (1976).
- Tjian, R. *Cell* **13**, 165-179 (1978).
- Shalloway, D., Kleinberger, T. & Livingston, D. *Cell* **20**, 411-422 (1980).
- Myers, R. & Tjian, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6491-6495 (1980).
- Van Heuverswyn, H. & Fiers, W. *Eur. J. Biochem.* **100**, 51-60 (1979).
- Scott, W. & Wigmore, D. *Cell* **15**, 1511-1518 (1978).
- Sundin, O. & Varshavsky, A. *J. molec. Biol.* **132**, 535-546 (1979).
- Varshavsky, A., Sundin, O. & Bohn, M. *Cell* **16**, 453-466 (1979).
- Saragosti, M., Moyne, G. & Yaniv, M. *Cell* **20**, 65-73 (1980).
- Wigmore, D., Eaton, R. & Scott, W. *Virology* **104**, 462-473 (1980).
- Stalder, J. *et al. Cell* **20**, 451-460 (1980).
- Wu, C. *Nature* **286**, 854-860 (1980).
- Flavell, R. *Nature* **285**, 356-357 (1980).
- Mathis, D., Oudet, P. & Chambon, P. *Prog. Nucleic Acids Res. molec. Biol.* **24**, 1-55 (1980).
- Mertz, J. & Berg, P. *Virology* **62**, 112-124 (1974).
- Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. *Cell* **14**, 725-731 (1978).
- Reddy, V. *et al. Science* **200**, 494-502 (1978).
- Legerski, R., Hodnett, J. & Gray, H. *Nucleic Acids Res.* **5**, 1445-1464 (1978).
- Fiers, W. *et al. Nature* **273**, 113-120 (1978).
- Sanger, F. & Coulson, A. *FEBS Lett.* **87**, 107-110 (1977).
- Favaloro, J., Treisman, R. & Kamen, R. *Meth. Enzym.* **65**, 718-749 (1980).

The SV40 early region TATA box is required for accurate *in vitro* initiation of transcription

Diane J. Mathis* & Pierre Chambon

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg, France

The SV40 early region TATA box functions in vitro to fix initiation of transcription within a narrow region identical to that in vivo. In the absence of the TATA box region initiation occurs at multiple specific sites, as in vivo. However, certain sequences far upstream, crucial for efficient in vivo expression, are not essential in vitro.

THE development of cell-free systems allowing selective and accurate initiation of transcription by RNA polymerase B (II)^{1,2} has given a new impetus to studies of the control of eukaryotic gene expression. Viral¹⁻³ and cellular^{3,4} genes have been transcribed using either a whole-cell² or cytoplasmic (S100)¹ extract

from HeLa or KB cells, and have generally shown close correspondence between the *in vitro* transcription start site for a given gene and its *in vivo* cap site (the nucleotide coding for the 5'-end of the mature mRNA).

This apparent fidelity has allowed the use of a HeLa cell-free system to delineate the DNA sequences required for specific RNA polymerase B initiation. Corden and co-workers⁵ have reported that the DNA segment 12-32 nucleotides upstream

* Present address: Stanford University School of Medicine, Department of Medical Microbiology, Stanford, California 95304, USA.

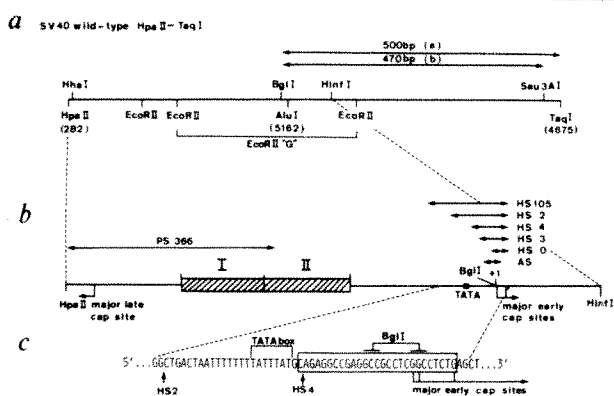


Fig. 1 Maps of the cloned 850-base pair SV40 *HpaII*-*TaqI* fragment. *a*, Relevant restriction enzyme sites. The SV40 sequence has been numbered according to the system proposed in ref. 12. Note that for our SV40 strain the *HpaII*-*TaqI* fragment lacks 8 base pairs which include an *EcoRII* site (nucleotides 108–116)¹³. The arrowed lines indicate the approximate size of expected run-off fragments (see Fig. 2). *b*, Location of mutant deletions and delineation of important functional elements. A more precise description of the mutants can be found in ref. 13. The major late cap site is that mapped by Ghosh and co-workers¹⁴; the major early cap sites were assigned according to refs 20 and 21, the accompanying paper¹³ and data within this paper. +1 corresponds to the early cap site at position 5,172. The stippled boxes represent the 72-base pair repeats. *c*, Sequence around the TATA box and wild-type cap sites²⁷. The arrows indicate the deletion end points for mutants HS2 and HS4. The boxed nucleotides are those included in a 27-base hairpin around the *Bgl*I site.

from the cap site of the adenovirus 2 (Ad-2) major late transcription unit suffices to promote specific initiation *in vitro*; similar results were obtained with a DNA segment 10–44 base pairs upstream from the conalbumin gene cap site⁵. These segments share the sequence 5'-CTATAAAAGGGG-3', which includes the TATA or 'Goldberg-Hogness' box (underlined) thought to be involved in promoting specific initiation, because it is conserved before the 5' end of most RNA polymerase B-transcribed genes^{6–9}. Moreover, specific RNA polymerase B initiation is essentially abolished on mutants with deletions involving the TATA sequence⁵, and when the conalbumin gene TATAAAA sequence is mutated to TAGAAAA, specific *in vitro* transcription is drastically reduced¹⁰. The conclusion, then, is that the TATA box has a fundamental role in promoting specific initiation by RNA polymerase B, at least *in vitro* using naked DNA templates and HeLa cell-free extracts.

Two studies cast doubt on whether the TATA box has an identical role *in vivo*. Deletion of the TATA box does not abolish transcription of the sea urchin H2A histone gene in *Xenopus laevis* oocytes¹¹ nor that of the SV40 early region in mammalian cells^{12,13}, but instead generates a new set of start sites. One might argue, of course, that these genes have atypical promoters—the histone genes are preceded by extensive A+T-rich ‘spacer’ regions and the SV40 early region TATA box is embedded in a highly A+T-rich region (5'-TAATTTTTTTT-TATTTAT-3'). Such sequence organizations are not typical of known RNA polymerase B-transcribed genes⁹, and the H2A genes and SV40 early region TATA boxes may not have retained their usual function(s). Alternatively, RNA polymerase B may not see the TATA box in the same context *in vivo* as *in vitro*.

It seems necessary, then, to assess the fidelity of the *in vitro* transcription systems at another level. Do mutations within a given DNA segment affect its *in vitro* and *in vivo* transcription identically? Here we compare the SV40 early region transcripts generated in a HeLa S100 extract with those detected *in vivo* in mammalian cells. We first establish that *in vitro* RNA polymerase B initiates transcription accurately on the early gene region, then compare the effects of various deletions in the promoter region on initiation *in vivo* and *in vitro*. The construction of these mutants and a detailed study of their *in vivo* expression are described in the accompanying article¹³.

RNA polymerase B can initiate accurately on the SV40 early region *in vitro*

Using a 'run-off' assay¹ we first determined whether RNA polymerase B could initiate accurately *in vitro* on the SV40 early region. A cloned fragment containing the SV40 early sequences *HpaII-TaqI* (see Fig. 1 and Fig. 2 legend) was purified and incubated with RNA polymerase B and ³²P-GTP in the presence of a HeLa S100 extract. The newly synthesized RNA was isolated, electrophoresed on a denaturing gel and detected by autoradiography. If initiation of transcription *in vitro* occurred at the major early start sites used *in vivo*, a 'run-off' RNA fragment about 500 base pairs long would be expected (Fig. 1; see below and ref. 13 for details of the assignment of major *in vivo* early cap sites). Transcripts initiating at the major *in vivo* late cap site¹⁴ would appear as short fragments of about 25 base pairs and should therefore not interfere with the assay. Two prominent 'run-off' bands were observed (Fig. 2, lane 2), one (*a*) about 500 base pairs long and possibly due to initiation at the major *in vivo* early start sites, the other (*b*) resulting from initiation about 80 base pairs downstream. We determined that

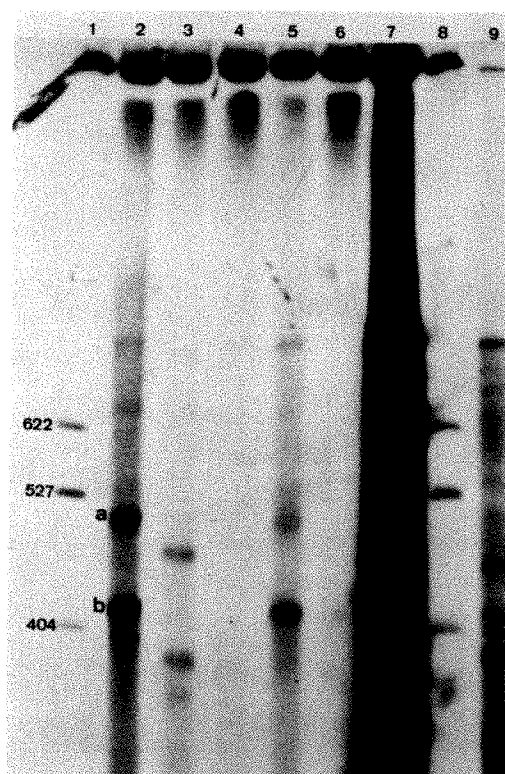


Fig. 2 Run-off analysis of RNA transcribed *in vitro* from wild-type SV40 early promoter. The 850-base pair *HpaII*-*TaqI* fragment from SV40 was inserted into the *EcoRI* site of pBR322 by blunt-end ligation after DNA polymerase I repair. The SV40 sequences could be excised from the resulting plasmid, pSV3, by *EcoRI* digestion, and were subsequently purified by sucrose gradient centrifugation. The purified fragment was incubated in a HeLa cell-free system and the run-off transcripts analysed as previously described^{3,5}, except for the following modifications. (1) The protocol for preparing S100 extract omitted the hypotonic cell washes but included a 30-min incubation of cells at 4°C in hypotonic buffer. Also, the cell homogenate was incubated at 0°C for 30–60 min after being rendered isotonic. (2) The NTP concentration during the RNA synthesis reaction was 200 µM ATP and UTP and 10 µM GTP and CTP. Lane 2, run-off transcripts from the 850-base pair *EcoRI* fragment from pSV3 (containing the Sv40 *HpaII*-*TaqI* sequence); lane 3, transcripts from the same *EcoRI* fragment subsequently cut by *Sau* 3A1; lane 4, 1 µg ml⁻¹ α-amanitin was included in the standard assay of transcripts from the pSV3 *EcoRI* fragment; lane 5, exogenous RNA polymerase B was omitted from the standard assay for transcripts from the pSV3 *EcoRI* fragment; lane 6, no DNA was added to the standard assay; lane 7, transcripts of the pSV3 *EcoRI* fragment generated in the absence of S100 extract; lane 9, lane 7 exposed for 1/15th the time; lanes 1 and 8, 5'-end-labelled *HpaII* digests of pBR322. Sizes for these sequenced marker fragments are indicated on the left.

these bands were due to synthesis from the early strand by using an *HpaII*-*Sau* 3AI fragment (Fig. 1a) as template in the *in vitro* system. We found that, as expected for transcripts synthesized from the early strand using this template, the two predominant RNA bands (a and b) obtained were about 30 base pairs shorter (Fig. 2, lane 3). That both bands represented transcription by RNA polymerase B was inferred from their absence when $1 \mu\text{g ml}^{-1}$ α -amanitin was included in the assay (Fig. 2, lane 4). When RNA polymerase B was omitted, the bands were present but at reduced intensity (Fig. 2, lane 5), indicating that active polymerase B exists in the S100 extract. No bands were detected when DNA was omitted (Fig. 2, lane 6), and in the absence of S100 extract the entire gel lane was blackened. On briefer exposure (lane 9) some bands were visible, but none corresponded to the two predominant start sites found in the presence of S100 extract. Thus, by the run-off assay, the *in vitro* system seems to allow specific initiation of RNA polymerase B at or near the major early cap sites used *in vivo* (a) and at a site about 80 base pairs downstream (b). Handa and co-workers¹⁵, using a HeLa whole-cell extract, have also detected initiation of transcription in these two regions of SV40.

S_1 nuclease mapping gave a more precise determination of the *in vitro* start sites corresponding to run-off band a. The early

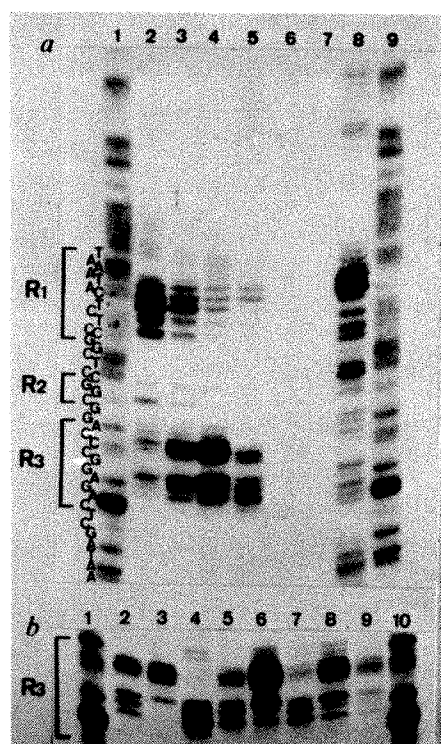


Fig. 3 S_1 nuclease mapping of the 5' ends of RNAs from infected CV1 cells or transcribed *in vitro* in the presence of the HeLa extract. Poly(A)⁺ cytoplasmic RNA was isolated from 30-h-infected CV1 cells as described in ref. 13. pSV3 transcripts were synthesized in HeLa extracts as in Fig. 2 (*HpaII*-*TaqI* template) and purified free of DNA as already described³. In all cases, 2.5-fold reactions were used for the mapping of *in vitro* transcripts. S_1 nuclease mapping, using a single-stranded *HpaII*-*HinfI* probe ³²P-5'-end labelled at the *HinfI* site (Fig. 1a), essentially was performed as in ref. 13. **a**, Lanes 2-4, fragments protected by 30-h-infected CV1 cell RNA after 3 h of treatment with 500 (lane 2), 1,000 (lane 3) or 2,000 (lane 4) U ml^{-1} of S_1 nuclease; lanes 5-8, fragments protected by *in vitro*-synthesized RNA after 3 h of treatment with 2,500 U ml^{-1} S_1 nuclease. The RNA synthesis reactions contained pSV3 *EcoRI* fragment and the S100 extract in standard conditions (lane 5), pSV3 *EcoRI* fragment and S100 in the presence of $1 \mu\text{g ml}^{-1}$ α -amanitin (lane 6), S100 in the absence of pSV3 template (lane 7) and pSV3 template in the absence of S100 (lane 8). Lanes 1 and 9 show an A+C sequence ladder of the SV40 *HpaII*-*HinfI* fragment²⁸. **b**, Lane 2, same as lane 4 of **a**; lanes 3-9, same as lane 5 of **a**, except that the NTP concentrations during the RNA synthesis reaction were 10 μM GTP and 200 μM CTP, ATP, UTP (lane 3), 10 μM CTP and 200 μM GTP, ATP, UTP (lane 4), 10 μM GTP, CTP and 200 μM ATP, UTP (lane 5), 10 μM ATP and 200 μM GTP, CTP, UTP (lane 6), 10 μM UTP and 200 μM ATP, GTP, CTP (lane 7), 10 μM UTP, ATP and 200 μM GTP, CTP (lane 8), 200 μM GTP, ATP, CTP and UTP (lane 9). Lanes 1 and 10 show an A+C sequence ladder of the SV40 *HpaII*-*HinfI* fragment²⁸.

coding strand of an *HpaII*-*HinfI* fragment (Fig. 1a), ³²P-end labelled at the *HinfI* site, was hybridized with either poly(A)⁺ cytoplasmic RNA from SV40-infected CV1 cells or RNA transcribed *in vitro* from the *HpaII*-*TaqI* fragment. The hybrids were treated with nuclease S_1 , denatured and electrophoresed on a sequencing gel. Progressively stronger S_1 treatment (Fig. 3a, lanes 2-4) shows that the DNA fragments protected by RNA synthesized *in vivo* has a complex pattern consisting of three components: R1, a cluster of bands mapping between nucleotides 5,189 and 5,197; R2, two bands, at positions 5,175 and 5,177; R3, two clusters of bands mapping at nucleotides 5,164-5,166 and 5,169-5,173. The bands in regions R1 and R2 decrease in intensity with stronger S_1 treatment, while the R3 bands increase. Most of the R1 and R2 bands map within a 27-base hairpin centred at the *BglII* site (Fig. 1c) and we suggest that the transient bands result from S_1 'pauses' in regions of extensive DNA secondary structure (this interpretation has been proposed to explain a similar phenomenon arising during S_1 mapping of conalbumin gene transcripts³). With this assumption we conclude that the major *in vivo* early cap sites fall within the region 5,164-5,166 and 5,169-5,173. S_1 mapping does not allow us to assign single-base start sites because S_1 trimming (according to the severity of digestion conditions) can result in 3'-ragged DNA with an overhang of 1-5 nucleotides¹⁶⁻¹⁹. Nevertheless, our assignment of the major *in vivo* early cap sites agrees with previously published results using reverse transcriptase primer extension mapping²⁰ or enzymatic cap analysis^{13,21}.

S_1 mapping of the *in vitro*-synthesized RNA results in an essentially identical pattern of protected fragments (Fig. 3a, lane 5), suggesting that RNA polymerase B initiates faithfully at the major *in vivo* cap sites, although a low amount of *in vitro* initiation may occur in regions 1 and 2. These S_1 -protected fragments were not observed when $1 \mu\text{g ml}^{-1}$ α -amanitin was included in the RNA synthesis reaction (Fig. 3a, lane 6), and they require the presence of DNA in the S100 extract (Fig. 3a, lane 7). In the absence of S100 extract (Fig. 3a, lane 8), initiation occurs at multiple sites throughout the fragment.

During these studies we noted a strange phenomenon related to the concentration of nucleotide triphosphates (NTPs) included in the *in vitro* RNA synthesis reaction. As described above, the major start sites *in vitro* consist of two clusters of bands (R3); however, this is not true if the concentration of purines in the synthesis reaction greatly exceeds that of pyrimidines or vice versa. For example, when GTP (or ATP) is present at 10 μM and the other three NTPs at 200 μM , the upper cluster of bands in R3 is much more prominent (Fig. 3b, lanes 3, 6). In contrast, when CTP (or UTP) is included at 10 μM and the other three NTPs at 200 μM , the lower cluster predominates (Fig. 3b, lanes 4, 7). Only if the total purine concentration approximates the total pyrimidine are both clusters of bands observed in a ratio similar to that found with *in vivo* RNA (Fig. 3b, compare lane 2 with lanes 5, 8 and 9). Although we cannot explain this phenomenon, it does suggest that our S_1 mapping detects true RNA polymerase start regions, as it is unlikely that the NTP concentration affects anything other than the *in vitro* RNA synthesis reaction. Also, any *in vitro* study of mutant promoter regions must consider this NTP concentration effect.

S_1 mapping of the RNA represented by run-off fragment b using the *EcoRII* G fragment (Fig. 1a) (data not shown) shows that the 5' end of the b transcript maps around nucleotide 5,080. Bands were not detected at this site in S_1 mapping of *in vivo* wild-type RNA even on heavily overloaded gels (although a band at 5,085 can sometimes be observed; data not shown). A 5'-GATAAA-3' sequence occurs 15 base pairs upstream from position 5,080, and a hairpin spans nucleotides 5,100-5,120 (centred ~30 base pairs upstream from the start site). We do not know whether either of these sequence elements corresponds to an initiation signal recognized *in vitro* but not *in vivo*.

Thus, we have shown that RNA polymerase B initiation on the SV40 early region occurs *in vitro* at a limited number of sites,

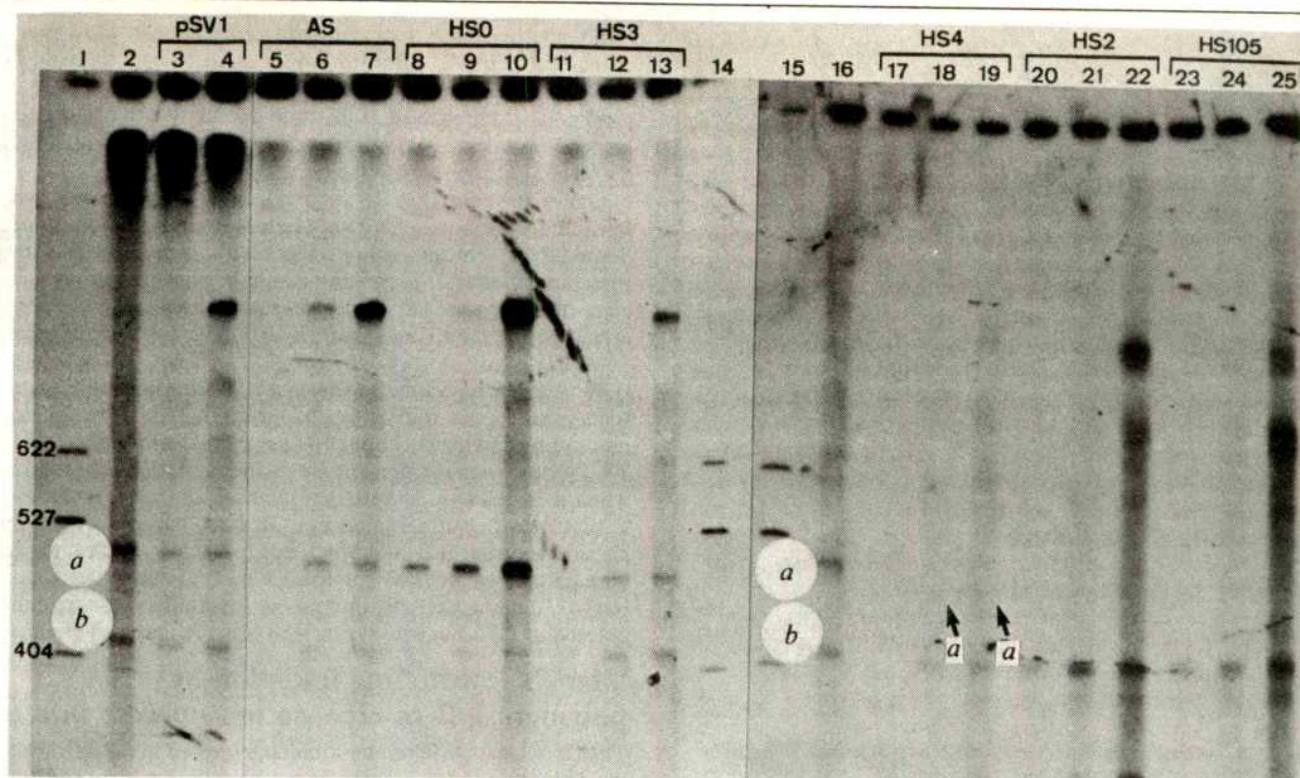


Fig. 4 Run-off analysis of RNA transcribed *in vitro* from deletion mutants in the promoter region. The construction of the wild-type plasmid pSV1 and the derivation of plasmids bearing deletions in the promoter region are described in ref. 13. Each of these recombinants was cut by *TaqI* and *HpaII* to provide an SV40 early region truncated template. The fragments were purified on sucrose gradients and then incubated for RNA synthesis in the presence of the HeLa S100 extract. The incubation conditions and subsequent analysis of transcripts were exactly as described in Fig. 2 legend. The RNA synthesis reactions contained for lanes 2 and 16: $4 \mu\text{g ml}^{-1}$ of the 850-base pair *EcoRI* fragment from pSV3; for lanes 3 and 4, 4 and $8 \mu\text{g ml}^{-1}$, respectively, of the *HpaII-TaqI* fragment from pSV1; for the remaining lanes: 2, 4 or $8 \mu\text{g ml}^{-1}$ of the *HpaII-TaqI* fragment from AS (lanes 5–7), HS0 (lanes 8–10), HS3 (lanes 11–13), HS4 (lanes 17–19), HS2 (lanes 20–22) or HS105 (lanes 23–25). The arrows in lanes 18 and 19 indicate the position of the band *a* transcripts visible on the original autoradiogram. All samples were derived from the same experiment although they were electrophoresed on two gels (lanes 1–14 and 15–25). The markers in lanes 1, 14 and 15 were ^{32}P -5'-end labelled *HpaII* fragments from pBR322.

one prominent set corresponding to the major cap sites used *in vivo*. Transcription seems rather inefficient, however, being about 20% that of the Ad-2 major late transcription unit in the same conditions (data not shown). Nevertheless, the fidelity of initiation with the HeLa cell-free system allows us to approach the question of whether mutations in the promoter region affect *in vitro* and *in vivo* initiation of transcription identically.

Some *in vivo* transcription signals are recognized *in vitro* by RNA polymerase B

The accompanying article¹³ describes the construction of a series of deletion mutants spanning the promoter region for the SV40 early genes. Deletions were of three types (Fig. 1b): (1) those removing sequences between the TATA box and the major *in vivo* cap sites (for example, AS), (2) those which start at position 5,162, delete the major cap sites and extend upstream towards and eventually beyond the TATA box (the HS series), and (3) those starting from the *HpaII* site at position 282 and extending downstream towards the TATA box (the PS series). RNA synthesized *in vivo* by the first two classes of mutants indicated that one function of the TATA sequence was to fix initiation of transcription within a narrow region¹³. In its presence, RNA polymerase B starts transcription about 30 base pairs downstream; in its absence, the enzyme initiates at multiple sites over a broad region. We now assess the effect of these promoter region deletions on *in vitro* RNA synthesis.

HpaII-TaqI fragments purified from the wild-type 'parental' plasmid pSV1 and from the various mutants, were transcribed *in vitro* and the run-off RNA transcripts analysed as described above (Fig. 2) for wild-type plasmid pSV3. (Note that pSV3 contains the SV40 early sequences *HpaII-TaqI* inserted at the *EcoRI* site of pBR322, whereas pSV1 contains the *HpaII-BamHI* sequences inserted at the same site.) Incubation of the pSV1 fragment in the cell-free system resulted in the expected

500 (*a*) and 420 (*b*) base pair run-off bands (Fig. 4, lanes 3, 4). Some bands not observed with the other wild-type plasmid—pSV3—were also evident (compare lanes 2 and 4), but these are not α -amanitin sensitive (data not shown) and probably correspond to full-length transcripts of contaminating DNA fragments [*HpaII-TaqI* sequences are easier to purify from pSV3 than from pSV1 or the various deletion mutants, especially when large deletions are involved (see Figs 2 and 4 legends for details)].

Mutants AS, HS0, HS3 and HS4 all retain the TATA box, although only AS retains the major *in vivo* cap sites (Fig. 1b). Run-off transcripts *a* and *b* are clearly produced when AS (Fig. 5, lanes 5–7), HS0 (lanes 8–10) and HS3 (lanes 11–13) fragments are incubated in the cell-free system. With HS4 (lanes 17–19), the *a* and *b* bands are both detectable, at least on the original autoradiogram, but *a* is much reduced. Closer inspection of the data reveals that with deletions which do not involve the TATA box (1) the size of band *a* decreases progressively with larger deletions, and (2) the ratio of *a* to *b* is different with the various mutants. Densitometric scanning of the autoradiograms indicates the following *a/b* ratios: pSV1 or pSV3, 1.0; AS, 2.5; HS0, 6.5; HS3, 1.0; HS4, 0.25. This effect has been observed in several experiments using different DNA preparations and varying NTP concentrations.

Mutants HS2 and HS105, which lack both the TATA box and the major *in vivo* cap sites (Fig. 1b), do not express an *a* transcript (lanes 20–22 and 23–25), although band *b* is readily apparent (a result obtained with several different DNA preparations and in all NTP conditions studied).

Using *S*₁ nuclease mapping to compare the effect of promoter region deletions on initiation *in vitro* and *in vivo* (Fig. 5a), we found that for those mutants which retain the TATA box (AS, HS0, HS3 and HS4), the major *in vivo* start sites are mimicked quite faithfully *in vitro*. In both environments, RNA polymerase B initiates about 30 base pairs downstream from the TATA box,

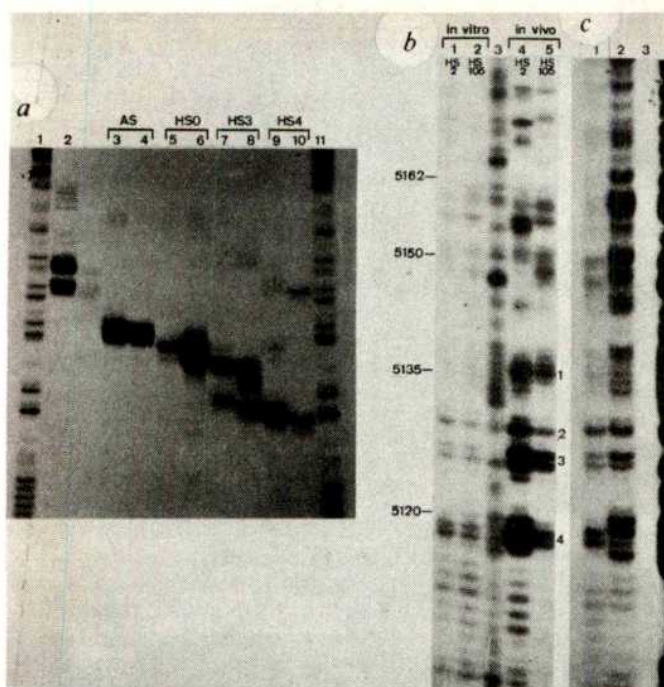


Fig. 5 S_1 nuclease mapping of 5' ends of RNAs synthesized *in vivo* or *in vitro* from mutant plasmids. Poly(A)⁺ cytoplasmic RNA was isolated from rat 3T3 cells transformed by the various mutants as described in ref. 13. S_1 mapping was as in Fig. 3 except that *HpaII-HinI* probes specific for each mutant were used. **a**, The cellular RNAs were from cell lines transformed by pSV1 (lane 2), AS (lane 3), HS0 (lane 5), HS3 (lane 7) or HS4 (lane 9). The *in vitro* RNA synthesis reactions contained *HpaII-TaqI* fragments from plasmids AS (lane 4), HS0 (lane 6), HS3 (lane 8) and HS4 (lane 10). Lanes 1 and 11 show an A+C sequence ladder²⁸ from the wild-type *HpaII-HinI* fragment. **b**, The cellular RNAs were from cell lines transformed by HS2 (lane 4) and HS105 (lane 5). The *in vitro* RNA synthesis reactions contained *HpaII-HhaI-TaqI* fragments from plasmids HS2 (lane 1) and HS105 (lane 2) (for mutants with long deletions, *HhaI* digestion was performed to facilitate separation of the SV40 fragment from contaminating pBR322 sequences). Lane 3 shows an A+C sequence ladder²⁸ from the wild-type *HpaII-HinI* fragment. The numbers to the left refer to nucleotide sequence number as in Fig. 1a; the numbers to the right correspond to band numbers in Fig. 3b of ref. 13. **c**, The *in vitro* RNA synthesis reactions contained *HpaII-HhaI-TaqI* fragments from HS105 in standard conditions (lane 1) in the absence of S100 (lane 2) or in the presence of $1 \mu\text{g ml}^{-1}$ α -amanitin (lane 3). For lane 2, only 1/8 of the S_1 -protected material was loaded.

regardless of the exact downstream sequences (ref. 13 gives a precise map of the *in vivo* start sites for these mutants). There are some differences in the RNAs synthesized *in vivo* and *in vitro*: for HS0, the ratio of the two starts differs in the two environments; for HS3 the *in vitro* start sites appear to extend further downstream than those of *in vivo* RNA. The experiments depicted in Fig. 5 give no indication of the efficiencies of transcription, as DNA probes of different specific activity were used for each mutant. However, S_1 -mapping of the mutant transcripts using a single wild-type probe (data not shown) shows that the level of initiation just downstream from the early region TATA box varied for different mutants, in the order HS0 > AS > HS3 $\approx$ PSV1 > HS4.

When the TATA box is deleted, early gene expression can still be detected *in vivo*, but transcription begins at multiple sites over a broad area¹³, and S_1 mapping experiments (Fig. 5b, c) demonstrate that this phenomenon also occurs *in vitro*. (This multiplicity of initiation sites probably explains why no distinct bands can be detected by the run-off assay; see above.) For a given mutant, the pattern of fragments protected by *in vitro*-synthesized RNA closely resembles that of fragments protected by *in vivo* transcripts (compare lanes 1 and 4 or lanes 2 and 5; see also Fig. 3b in ref. 13). Most bands are observed with both *in vivo* and *in vitro* RNAs, although relative band intensities differ, sometimes markedly. Using a single wild-type probe we also determined that the *in vitro* transcription of HS2 is not much less than that of the wild-type plasmid pSV1 (data not shown),

although the new set of cap sites generated in HS2 was not detected for pSV1, indicating that, as *in vivo*¹³, the wild-type TATA sequence may 'mask' its functional substitutes. We do not know whether these 'substitute' elements are preferentially located in the early promoter region or scattered throughout the SV40 genome.

Figure 5c shows that RNA polymerase B alone selects specific initiation sites, but yields a very heterogeneous pattern (lane 2). Some factor(s) of the S100 extract restricts initiation to a few of these sites (lane 1), producing a similar pattern to that seen *in vivo*. It also requires RNA polymerase B, being abolished by the inclusion of $1 \mu\text{g ml}^{-1}$ α -amanitin in the RNA synthesis reaction (lane 3).

We therefore conclude that some of the initiation 'signals' recognized *in vivo* by RNA polymerase B are also recognized *in vitro*. Initiation occurs about 30 base pairs downstream from the TATA box when it is present, but at many sites over a broad area when it is deleted. The 'fixing' function probably includes the TATA box because initiation ~ 30 base pairs downstream is already reduced in HS4, which has a deletion end point just 2 base pairs downstream from the TATTTAT sequence (Fig. 1c), but we cannot exclude the possibility that other sequences, deleted in HS2 but present in HS4 (Fig. 1c), are involved.

'Signals' controlling initiation by RNA polymerase B *in vivo* do not operate *in vitro*

The third class of deletion mutants described in the accompanying article¹³, the PS series, retains the major wild-type cap sites as well as a normal TATA box, but lacks sequences further upstream. Analysis of the *in vivo* expression of PS mutants

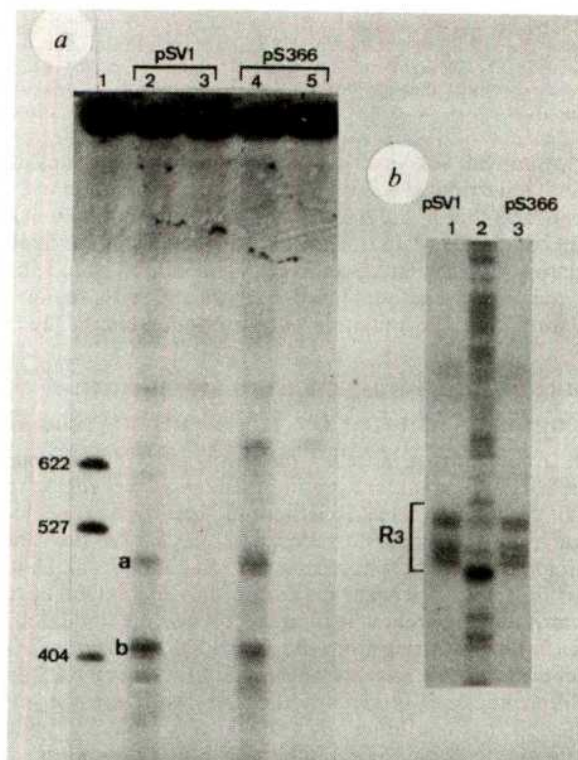


Fig. 6 Run-off analysis and S_1 nuclease mapping of RNA synthesized *in vitro* from the deletion mutant PS366. The construction of PS366 is described in ref. 13. **a**, Run-off analysis was as in Fig. 2. Lane 2, the *HpaII-TaqI* fragment from pSV1 was incubated with S100 extract in standard conditions. Lane 3, the same fragment incubated with S100 in the presence of $1 \mu\text{g ml}^{-1}$ α -amanitin. Lane 4, the *HpaII-HhaI-TaqI* fragment from PS366 was incubated with S100 in standard conditions. Lane 5, the same fragment was incubated with S100 in the presence of $1 \mu\text{g ml}^{-1}$ α -amanitin. Lane 1 shows a ^{32}P -5'-end-labelled *HpaII* digest of pBR322. **b**, S_1 nuclease mapping was as in Fig. 5. Lane 1, fragments protected by RNA synthesized *in vitro* from the pSV1 template; lane 2, an A+C sequence ladder²⁸ of the wild-type *HpaII-HinI* fragment; lane 3, fragments protected by RNA synthesized *in vitro* from the PS366 template. R3 refers to the region delineated in Fig. 3a.

indicates that sequences mapping in the region of the two 72-base pair repeats I and II (Fig. 1b) markedly affect early gene expression. For example, PS366, which lacks nucleotides 98–282 and thus more than one complete copy of the 72-base pair repeat (Fig. 1b), does not detectably express the early genes during short-term infection of CV1 cells, and transforms rat 3T3 fibroblasts much less efficiently than a wild-type plasmid. However, Fig. 6a demonstrates that, by the run-off assay, *in vitro* transcription of PS366 is not similarly inefficient. The *a* transcript is readily apparent when an *HpaII*–*TaqI* fragment from PS366 is transcribed in the HeLa cell-free system (lane 4). Furthermore, S₁ nuclease mapping reveals that RNA polymerase B initiates the *a* transcript at the same sites on PS366 as on the wild-type pSV1 (Fig. 6b, lanes 1, 3). Thus, those sequence elements in the region of the 72-base pair repeats that exert a crucial influence on *in vivo* expression have no apparent effect in an *in vitro* environment.

***In vitro* and *in vivo* promoter regions for RNA polymerase B**

The three DNA sequence elements implicated in the control of initiation of transcription by RNA polymerase B are the TATA box region^{3,5,10,11,13}, the cap site region^{5,11} and a region upstream from the TATA box (refs 11–13, and personal communications from R. Grosschedl and M. L. Birnstiel, and from R. Kamen).

The present report and the accompanying one suggest that, both *in vivo* and *in vitro*, the early gene TATA box (or closely adjacent sequences) functions to limit initiation to a narrow region. In its presence, RNA polymerase B starts transcription about 30 base pairs downstream, and a similar function has been noted for other genes^{5,11,13,22}. Whatever actually measures the TATA box–cap site distance operates quite accurately *in vitro*, even in a heterologous system, for the newly selected cap sites on mutant templates are similar to those used in mammalian cells. The finding that on deletion of the early gene TATA box RNA polymerase initiates at multiple sites over a broad region both *in vitro* and *in vivo* is in keeping with a report that, in *Xenopus* oocytes, initiation occurs at multiple sites on an injected sea urchin histone gene lacking its normal TATA sequence¹¹. The multiple initiation sites we observe are nearly identical to those found *in vivo* (Fig. 5b), and this specific pattern is dependent on the presence of some factor(s) present in the cell-free extract (Fig. 5c), suggesting that the factor(s) which recognize the TATA box could also recognize its functional substitutes. Deletions or single-point mutations involving the TATA box have been reported to abolish, or drastically reduce, *in vitro* transcription^{5,10}, but such studies relied mainly on the run-off assay, and we also were unable to detect transcripts from the TATA-minus mutants using this assay. Alternatively, it is possible that some genes, like the SV40 early region, have sequences which can substitute functionally for the TATA box, while other genes either do not have them or have sequences which function much less efficiently than the TATA sequence.

We have observed that in mutants retaining the TATA box, the newly selected cap sites seem to be used with varying efficiencies *in vitro*, and this has been attributed^{5,11} to control by sequences around the cap site. We cannot decide whether the crucial sequences are those at the cap site and/or those between the TATA box and start site. Clearly, the mere presence of purines around the initiating nucleotide is not the important factor, because initiation on both HS0 and HS4 occurs in a region rich in purines. (*In vivo* transcription of many genes by RNA polymerase B starts with an A within a pyrimidine-rich region^{5,9}.) Nucleotides between the TATA box and the start sites might well influence the efficiency of cap site utilization, especially for the SV40 early genes, where a 27-base hairpin normally separates the TATA and the cap site (Fig. 1c). It is possible that AS and HS0 are transcribed so efficiently merely because this hairpin is disrupted, and that HS4 is transcribed much less efficiently because its longer deletion begins to affect the TATA box function. Interestingly, similar 'up'-mutations have been noted *in vivo*. Rat cells transformed by plasmids

bearing deletions 10–15 base pairs downstream from the early gene TATA box produce more early mRNA than cells transformed by wild-type plasmid (see ref. 22 and the accompanying article). Gluzman *et al.*²² have suggested that this overproduction of early mRNA might be due to a failure in autoregulation by T antigen. In light of the *in vitro* results, it might, instead or in addition, be caused by a more efficient utilization of the newly selected cap sites.

The far upstream sequences¹³, which are required *in vivo* but not *in vitro*, might bind RNA polymerase directly (analogous in some respects to the prokaryotic 'recognition site'²³), might dictate chromatin structure directly²⁴ or indirectly, or might bind activator proteins. We have shown that the mutant plasmid PS366, which lacks certain sequences upstream from the TATA box and is expressed inefficiently *in vivo*, is transcribed quite readily in the HeLa cell-free extract. The level of initiation approximates that for wild-type plasmid and the start sites are those normally used. This observation tends to argue against a model whereby the upstream sequences exert their influence simply by binding polymerase. However, the efficiency of transcription *in vitro* is much lower than *in vivo*^{1–5}, and it is possible that the mode of polymerase binding differs in the two environments. It may be that the chromatin structure in the promoter region brings into close vicinity the TATA box and the upstream sequences and that accurate reconstruction of chromatin structure might be important in understanding the role of the upstream sequences. We do not know the state of the DNA after incubation in the S100 extract, but it is unlikely to be packaged in chromatin as it is *in vivo*, for we have observed that during the standard RNA synthesis reaction supercoiled DNA templates are converted to relaxed covalently closed circles (unpublished results), whereas if nucleosomes had been formed, supercoiled DNA should have been recovered^{23,24}. Certainly, activator proteins could also be missing from the cell-free system. Clues as to how the upstream sequences control initiation *in vivo* and why they do not exert a similar influence *in vitro* await further experimentation.

We thank Drs C. Benoist, B. Davison and C. K  dinger for the gift of materials and for helpful advice, Mr J. M. Garnier for technical assistance, and Ms B. Chambon, C. Kutschis and Mr B. Boulay for help in preparing the manuscript. This work was supported by grants from the INSERM (ATP 72.79.104, CL 801029 and CL 801033), the CNRS (ATP 'Biologie Mol  culaire de G  ne'), the Fondation Fran  aise pour la Recherche M  dicale and from the Fondation S. and C. Del Duca.

Received 1 October 1980; accepted 22 January 1981.

- Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. *Cell* **18**, 469–484 (1979).
- Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gelfand, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3855–3859 (1980).
- Waslyk, B., K  dinger, C., Corden, J., Brison, O. & Chambon, P. *Nature* **285**, 367–373 (1980).
- Luse, D. S. & Roeder, R. G. *Cell* **20**, 691–699 (1980).
- Corden, J. *et al. Science* **209**, 1406–1414 (1980).
- Gannon, F. *et al. Nature* **278**, 428–434 (1979).
- Proudfoot, N. J. *Nature* **279**, 376 (1979).
- Flavell, R. A. *Nature* **285**, 356–357 (1980).
- Breathnach, R. & Chambon, P. A. *Rev. Biochem.* (in the press).
- Waslyk, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7024–7028 (1980).
- Grosschedl, R. & Birnstiel, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1432–1436 (1980).
- Benoist, C. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3865–3869 (1980).
- Benoist, C. & Chambon, P. *Nature*, **290**, 304–310 (1981).
- Ghosh, P., Reddy, V., Swinscoe, J., Lebowitz, P. & Weissman, S. *J. molec. Biol.* **126**, 813–846 (1979).
- Handa, H., Kaufman, R. J., Manley, J., Gelfand, M. & Sharp, P. A. *J. biol. Chem.* (in the press).
- Hentschel, C., Irminger, J. C., Bucher, P. & Birnstiel, M. L. *Nature* **285**, 147–151 (1980).
- Weaver, R. F. & Weissman, C. *Nucleic Acids Res.* **7**, 1175–1195 (1979).
- Sollner-Webb, B. & Roeder, R. *Cell* **18**, 485–499 (1979).
- Tsujimoto, Y. & Suzuki, Y. *Cell* **16**, 425–436 (1979).
- Reddy, V. B., Ghosh, P. K., Lebowitz, P., Piatak, M. & Weissman, S. M. *J. Virol.* **30**, 279–296 (1979).
- Haegeman, G. & Fiers, W. *J. Virol.* **35**, 955–961 (1980).
- Gluzman, Y., Sambrook, J. F. & Frisque, R. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3898–3902 (1980).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319–353 (1979).
- Waslyk, B., Oudet, P. & Chambon, P. *Nucleic Acids Res.* **7**, 705–713 (1979).
- Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843–1847 (1975).
- Birkenmeier, E. H., Radonovich, M. F., Shani, M. & Salzman, N. P. *Cell* **11**, 495–504 (1977).
- Van Heuverswyn, H. & Fiers, W. *Eur. J. Biochem.* **100**, 51–60 (1979).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499–580 (1980).

LETTERS

Pulsar and diffuse contributions to observed galactic γ radiation

A. K. Harding & F. W. Stecker

Laboratory for High Energy Astrophysics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

Satellite data on the energy spectrum of galactic γ -radiation^{1,2}, shows that such radiation has a multicomponent nature³⁻⁷. Here we use a calculation of the pulsar γ -ray emission spectrum together with a statistical analysis of recent data on 328 known pulsars to make a new determination of the pulsar contribution to galactic γ -ray emission. We then re-examine the contributions from diffuse interstellar cosmic-ray induced production mechanisms to the total emission and conclude that pulsars may account for a significant fraction of galactic γ -ray emission.

Models for the production of high-energy γ rays in the Galaxy have dealt primarily with diffuse emission processes, those which involve interactions between cosmic rays and interstellar matter. The major contributions are thought to come from electron bremsstrahlung and the decay of neutral pions produced in high energy proton collisions, with a small additional flux from Compton interactions. In determining the relative contributions of these processes from the shape of the observed γ -ray spectrum, the contribution of true point sources (as opposed to enhancements in the spatial distribution from diffuse processes) has been neglected, largely because the point source spectrum has been undetermined. At present, pulsars can give the best information on a true point source component of the galactic γ -ray flux.

We report here on the first calculation of a γ -ray production spectrum from pulsars in the Galaxy and on the implications of this point source contribution to the general interpretation of the observed galactic γ -ray spectrum. The details of this calculation, as well as a determination of the longitude and latitude profiles of pulsar γ rays will be described elsewhere⁸. There have been previous estimates of the pulsar contribution to the observed galactic γ -ray flux⁹⁻¹¹, using only 51 pulsars and very general assumptions about their γ -ray luminosities. There are now 328 known radio pulsars, all having independent distance determinations through their dispersion measures. The local density and galactic distribution can therefore be much better determined, assuming a value for the mean free electron density in the galactic plane. Pulsed γ rays have been observed from the Crab (PSR0531+21) and Vela (PSR0833-45) pulsars by the SAS 2 satellite^{12,13}, and the COS B satellite¹⁴. The energy spectra of these sources have recently been determined in the range 50 MeV–2 GeV and they can both be fit with steep power laws of similar slope around -2 (ref. 15). Although these two young sources represent only a small fraction of the known pulsars, their emission testifies to the fact that some, and possibly all pulsars are γ -ray emitters at some stage of their evolution. Though these others may be individually below the present limits of detectability, the contribution of their γ rays to the observed galactic emission could be significant, as has been previously suggested⁹.

We assume that the γ -ray emission above 100 MeV is produced by curvature radiation from primary particles which have been accelerated to high energies in the electric fields of the pulsar's magnetosphere. Some of these γ rays are converted to electron-positron pairs in the strong pulsar magnetic field. The electric field induced perpendicular to the magnetic field by the star's rotation also contributes to the pair production rate and has a large effect on the optical depths in short period pulsars¹⁶. The shape of the numerically calculated spectra in this model

depend on the initial energy of the primary particles and on the pulsar magnetic field strength and period¹⁷. For an initial energy of 1.5×10^4 GeV and a field strength of 10^{12} G, the resulting calculated spectrum for a pulsar period of 0.033 s reproduces the shape of the observed Crab pulsar spectrum. The observed 100-MeV flux from the Crab pulsar at 2 kpc is used to determine the normalization factor ($\sim 10^3$) which is the ratio of the primary particle density above the polar cap to the corotation charge density. Using this same normalization factor, primary particle energy, and magnetic field strength, the calculated spectrum for a period of 0.089 s fits the observed Vela pulsar spectrum quite well. It may be reasonable to assume, then, that these values of the parameters determining the spectrum are valid for all pulsar periods.

An expression for the γ -ray luminosity can be derived by integrating the spectra calculated for different periods, P , and surface magnetic field strengths, B_{12} (in units of 10^{12} G). The best fit for the dependence of luminosity on these quantities gives:

$$L_{\gamma}(>100 \text{ MeV}) = 1.2 \times 10^{35} B_{12}^{0.95} P^{-1.7} \text{ photons s}^{-1} \quad (1)$$

The γ -ray luminosities and fluxes for the known radio pulsars have been calculated from equation (1) and compared with the upper limits obtained by SAS 2 (ref. 18) and the sources in the COS B catalogue¹⁹. Eight pulsars have predicted fluxes above the point source detectability limit of COS B but most are towards the galactic centre where the diffuse background is high. Only two lie in the longitude range $90^\circ < l < 270^\circ$ where the search is considered complete down to $1.3 \times 10^{-6} \text{ cm}^{-2} \text{ s}^{-1}$; however, they both lie just outside the latitude range of the search. Four pulsars have predicted fluxes above their SAS 2 upper limits, but exceed them by no more than the estimated uncertainties in either the measurements or the model.

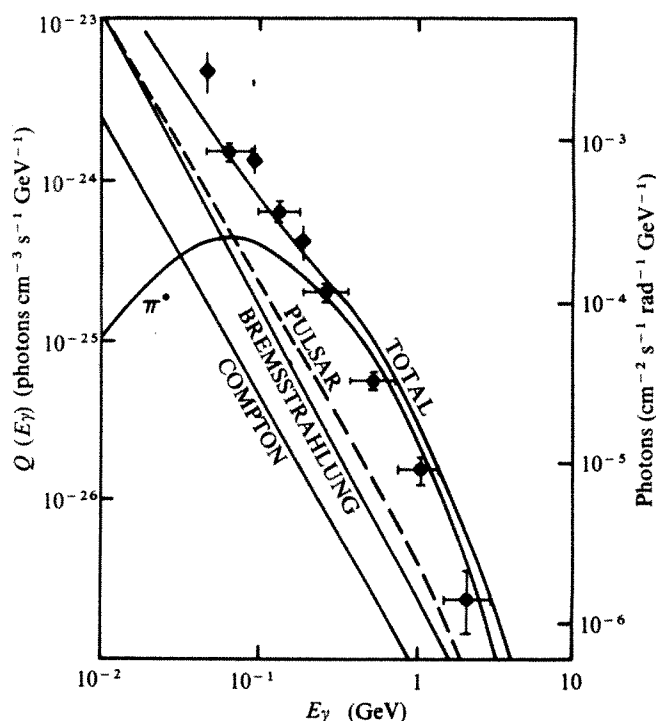


Fig. 1 Differential production spectra for pulsars and the major diffuse processes. The π^0 , bremsstrahlung and Compton spectra are from refs 6, 26 and the total spectrum includes the pulsar contribution. SAS 2 and COS B data points for the longitude range $355^\circ < l < 15^\circ$ are also shown ($\diamond$, ref. 2; $\bullet$, ref. 1). The left- and right-hand scales refer to the theoretical production rates at 10 kpc from the galactic centre and the observed γ -ray flux, respectively.

In calculating these pulsar spectra and luminosities, it has been assumed that the magnetic field is a pure dipole and that the angle between the dipole axis and spin axis is 90° , because we are concerned here with the collective properties of the pulsar population, and a random distribution of angles would give an average of 90° . More details of the model are given in ref. 17.

If all pulsars produce pulsed γ rays according to this model, then the local differential production spectrum will be,

$$q_{\text{PSR}}(E_\gamma) = \frac{D_\odot}{2h} \int_0^\infty I_\gamma(P, E_\gamma) \rho(P) dP$$

where $D_\odot$ is the local surface density of pulsars which are beamed in our direction, h is their scale height above the plane, $\rho(P)$ is their period distribution and $I_\gamma(P, E_\gamma)$ is the γ -ray spectrum of an individual pulsar with period P . $I_\gamma(P, E_\gamma)$ is also directly proportional to surface magnetic field strength, so that q_{PSR} will be proportional to the average pulsar magnetic field strength. We take an average field of 10^{12} G, which is consistent with measured pulsar spindown rates. The derived local density and scale height depend on a determination of $\langle n_e \rangle$, the mean electron density in the galactic plane, which, through dispersion measure observations, fixes the pulsar distance scale. Unfortunately, the galactic distribution of n_e is not well known and is difficult to determine. There are several indications that n_e is not uniform throughout the galactic plane and that it does not have a simple exponential Z dependence^{20,22}, both of which have been previously assumed in determining pulsar densities. However, in view of the difficulties entailed in a detailed spatial modeling of n_e , given the present observational data, we adopt the simple constant density model of Taylor and Manchester²³ which is sufficient for estimating the mean contribution from pulsars to the galactic γ -radiation. For $\langle n_e \rangle = 0.03 \text{ cm}^{-3}$, a value consistent with H I absorption distances to pulsars in local regions of the Galaxy²⁰, a statistical analysis of data on the 328 known pulsars gives $D_\odot = 480 \text{ pulsars kpc}^{-2}$ and $h_p = 325 \text{ pc}$. The resulting γ -ray production spectrum per cm^3 is shown in Fig. 1 along with the production spectra for the major diffuse processes calculated⁶, assuming a local interstellar hydrogen density of 1 atom cm^{-3} . This spectrum gives volume-averaged integral pulsar γ -ray production rates at 10 kpc of 2.2×10^{-26} photons $\text{cm}^{-3} \text{ s}^{-1}$ above 100 MeV and 4.6×10^{-26} photons $\text{cm}^{-3} \text{ s}^{-1}$ above 35 MeV. Although the calculated pulsar spectrum has been extrapolated below 100 MeV, the model does not accurately predict the low energy spectrum because the synchrotron radiation from secondary particles, which may extend to the γ -ray region, has not been included. This contribution would tend to steepen the spectrum and increase the estimated flux level below around 100 MeV.

The diffuse production spectra in Fig. 1 have widely differing uncertainties. The pion production spectrum is least uncertain, because it primarily involves cosmic ray protons with energies in the range 1–10 GeV where uncertainties in determining galactic fluxes owing to solar modulation effects produce small uncertainties ($\leq 15\%$) in the total production rate²⁴. Even the differential γ -ray source spectrum for $E_\gamma \leq 10 \text{ GeV}$ is insensitive to the kinematical model chosen, as shown by the excellent agreement between calculations given in refs 24–28.

The Compton production mechanism produces only a small expected contribution in our region of the Galaxy, however, this process could be significant near the galactic centre⁶. As this component is small over most of the galaxy and various theoretical estimates^{29,30} are in reasonable agreement with Fig. 1, we conclude that little error in determining the total galactic flux is introduced by the Compton process.

The situation is quite different in the case of electron bremsstrahlung. The curve shown in Fig. 1 for this component was determined using a conservative 'standard' cosmic ray electron spectrum³¹. However, the electron spectrum cannot be easily determined from cosmic-ray data due to large solar modulation effects. Theorists have thus had much leeway in invoking large electron fluxes to account for discrepancies in both the spectrum and intensity of the observed galactic γ rays with 'standard'

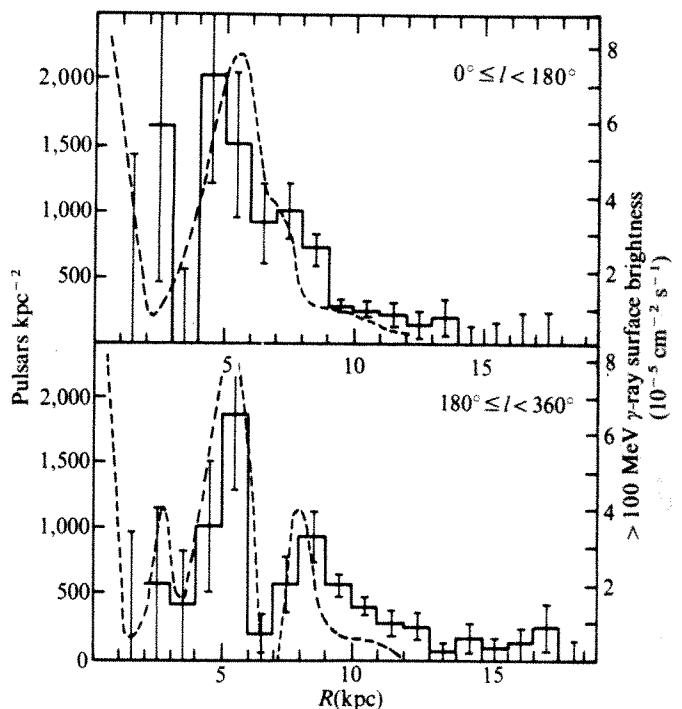


Fig. 2 Surface density of pulsars (solid line)⁸ and γ -ray surface emissivity (dashed line)⁶ plotted against galacto-centric radius R for northern and southern halves of the Galaxy.

predictions^{32,33}. Indeed, spectral data have been used to speculate on drastic fluctuations in both electron and proton fluxes in the Galaxy³⁴, although we believe that inherent systematic errors as well as statistical errors in the spectral data make such a radical analysis unwarranted.

The present calculated pulsar contribution has clear implications for analysing the total galactic spectrum. Pulsars contribute a soft power law flux component similar to that expected from electron bremsstrahlung. In addition, the pulsar component as estimated in Fig. 1 is as large as the 'standard' bremsstrahlung component.

We also expect that the pulsar contribution to the total galactic flux should be roughly independent of galactic longitude except in the outer Galaxy. We have used the longitude and distance data for the full sample of observed pulsars, together with corrections for selection effects, to determine the radial distributions of pulsars in both the northern ($0^\circ < l < 180^\circ$) and southern ($180^\circ < l < 360^\circ$) galactic longitudes. The results for the inner Galaxy ($R \leq 8 \text{ kpc}$) are strikingly similar to the distributions of total γ -ray emissivity determined previously⁶. Thus, the ratio of pulsar to total emissivity (the fraction of the flux from the pulsar component) should be roughly constant with longitude, for longitudes within 60° of the galactic centre. This fraction is calculated to be 15–20% above 100 MeV (as derived from both Fig. 1 and a comparison of the longitude distribution of pulsar and observed γ -ray emission⁸ using the data in Fig. 2). It can easily be of the order of 50% at lower energies (see Fig. 1), removing the need to invoke large low-energy electron fluxes to explain the shape of the observed γ -ray spectrum.

The latitude distribution of pulsar γ -ray emission will be narrow but will extend somewhat beyond $b = 10^\circ$ (ref. 8). Pulsars may therefore be able to account for some of the flux observed^{35,36} at galactic latitudes above 10° . According to the model, the younger pulsars are the strongest γ -emitters. They are also more confined to the galactic plane³⁷. Thus, the pulsars at high b have a higher mean age and are weaker. This would be consistent with the lack of observed point sources at high b .

We conclude therefore that the integrated flux from distant γ -ray emitting pulsars may provide an important fraction of the low- to medium-energy galactic γ -ray flux (although a smaller fraction of the flux above 100 MeV), and the inclusion of the

pulsar component in analysing the galactic γ -ray spectrum supports a coherent integrated relationship with galactic cosmic-ray theory as discussed elsewhere³⁸⁻⁴².

Received 16 October 1980; accepted 17 February 1981.

1. Paul, J. A. *et al.* *Astr. Astrophys.* **63**, L31-L33 (1978).
2. Hartman, R. C. *et al.* *Astrophys. J.* **230**, 597-606 (1979).
3. Stecker, F. W. *Cosmic Gamma Rays* (Mono, Baltimore, 1971).
4. Stecker, F. W., Puget, J. L., Strong, A. W. & Bredekamp, J. H. *Astrophys. J. Lett.* **188**, L59-L61 (1974).
5. Fichtel, C. E. *et al.* *Astrophys. J.* **208**, 211-219 (1976).
6. Stecker, F. W. *Astrophys. J.* **212**, 60-70 (1977).
7. Higdon, J. C. *Astrophys. J.* **232**, 113-128 (1979).
8. Harding, A. K. *Astrophys. J.* (in the press).
9. Higdon, J. C. & Lingenfelter, R. E. *Astrophys. J. Lett.* **208**, L107-L110 (1976).
10. Strong, A. W., Wolfendale, A. W. & Dahanayake, C. *Mon. Not. R. astr. Soc.* **179**, 69-72 (1977).
11. Hartman, R. C. *et al.* in *The Structure and Content of the Galaxy and Galactic Gamma Rays* (eds Fichtel, C. E. & Stecker, F. W.) 15-25 (NASA CP-002, U.S. Government Printing Office, Washington DC, 1977).
12. Kniffen, D. A. *et al.* *Nature* **251**, 397-398 (1974).
13. Thompson, D. J., Fichtel, C. E., Kniffen, D. A. & Ögelman, H. *Astrophys. J. Lett.* **200**, L79-L82 (1975).
14. Bennett, K. *et al.* *Astr. Astrophys.* **61**, 279-284 (1977).
15. Lichti, G. G. *et al.* *Non-Solar Gamma Rays* (COSPAR Symp.) (eds Cowsik, R. & Wills, R. D.) 49-53 (Pergamon, New York, 1980).
16. Harding, A. K., Tademaru, E. & Esposito, L. W. *Astrophys. J.* **225**, 226-236 (1978).
17. Harding, A. K. *Astrophys. J.* (in the press).
18. Ögelman, H., Fichtel, C. E., Kniffen, D. A. & Thompson, D. J. *Astrophys. J.* **209**, 584-591 (1976).
19. Swanenburg, B. N. *et al.* *Astrophys. J.* (in the press).
20. Weisberg, J. M., Rankin, J. & Boriakoff, V. *Astr. Astrophys.* **88**, 84-93 (1980).
21. Ables, J. G. & Manchester, R. N. *Astr. Astrophys.* **50**, 177-184 (1976).
22. Harding, A. K. *Pulsars* (eds Sieber, W. & Wiełebinski, R.) (Reidel, Dordrecht, in the press).
23. Taylor, J. H. & Manchester, R. N. *Astrophys. J.* **215**, 885-896 (1977).
24. Stecker, F. W. *Astrophys. J.* **185**, 499-504 (1973).
25. Stecker, F. W. *Astrophys. Space Sci.* **6**, 377-389 (1970).
26. Stecker, F. W. *Astrophys. J.* **228**, 919-927 (1979).
27. Ganguli, S. N. & Sreekantan, B. V. *J. Phys.* **A9**, 311-321 (1976).
28. Badhwar, G. D. & Stephens, S. A. *Proc. 15th Int. Cosmic Ray Conf. Plovdiv* **1**, 198-203 (1977).
29. Stecker, F. W. in *Origin of Cosmic Rays* (eds Osborne, J. L. & Wolfendale, A. W.) 267-334 (Reidel, Dordrecht, 1975).
30. Piccinotti, G. & Bignami, G. F. *Astr. Astrophys.* **52**, 69-75 (1976).
31. Daugherty, J. K., Hartman, R. C. & Schmidt, P. J. *Astrophys. J.* **198**, 493-505 (1975).
32. Cesarsky, D. J., Paul, J. A. & Shukla, P. G. *Astrophys. Space Sci.* **59**, 73-83 (1978).
33. Schlickeiser, R. & Thielheim, K. O. *Mon. Not. R. astr. Soc.* **182**, 103-109 (1978).
34. Strong, A. W., Wolfendale, A. W., Bennett, K. & Wills, R. D. *Mon. Not. R. astr. Soc.* **182**, 751-760 (1978).
35. Fichtel, C. E., Simpson, G. A. & Thompson, D. J. *Astrophys. J.* **222**, 833-849 (1978).
36. Lebrun, F. & Paul, J. A. *Proc. 16th Int. Conf. Cosmic Rays* 13-18 (1979).
37. Taylor, J. H. in *The Large-Scale Characteristics of the Galaxy* (ed. Burton, W. B.) 119-123 (Reidel, Dordrecht, 1979).
38. Stecker, F. W. *Phys. Rev. Lett.* **35**, 188-191 (1975).
39. Stecker, F. W. *Nature* **260**, 412-414 (1976).
40. Stecker, F. W. & Jones, F. C. *Astrophys. J.* **217**, 843-858 (1977).
41. Stecker, F. W. *Am. Sci.* **66**, 570-576 (1978).
42. Stecker, F. W. *The Large-Scale Characteristics of the Galaxy* (ed. Burton, W. B.) 475-782 (Reidel, Dordrecht, 1979).

Changes in orientation and appearance of the 100 m arc s radio structure in SS433

R. T. Schilizzi*, G. K. Miley†, J. D. Romney‡ & R. E. Spencer§

* Netherlands Foundation for Radio Astronomy, Radio Sterrenwacht, Dwingeloo, The Netherlands

† Sterrewacht Leiden, Leiden, The Netherlands

‡ Max-Planck-Institut für Radioastronomie, Bonn, FRG

§ University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

The discovery¹ of radio structure on a scale of 10–50 m arc s in SS433 has prompted us to monitor the object with the European VLBI network. The first observations in this programme were made for periods of 9 and 10 h at a wavelength of 6 cm on 9 January (JD2444248) and 1–2 June (JD2444392.5) 1980. Our results, together with the observations of Walker *et al.*² at an earlier epoch, show that the compact radio structure of SS433 changes orientation in a manner consistent with the 164-day precession period inferred from the optical emission lines, but with a phase delay of 5 ± 4 days. The shape of the radio structure is elongated, non-linear and changes from one observing epoch to the other.

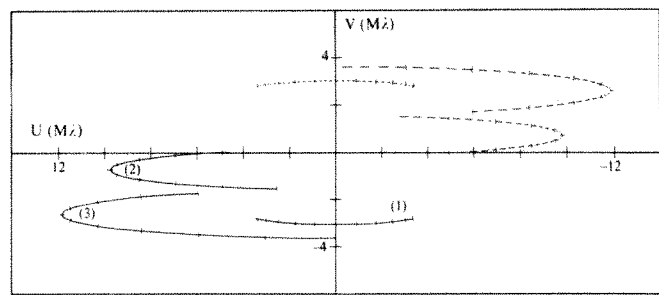


Fig. 1 Loci of the east-west (U) and north-south (V) components of the projected baselines for our 6-cm observations of SS433. The baselines shown are those for which the source was detected: (1) Westerbork-Effelsberg; (2) Chilbolton-Effelsberg; (3) Knockin-Effelsberg.

Signals were recorded with the standard MkII VLBI system (bandwidth 2 MHz) and the cross correlations were performed at the Max-Planck-Institut für Radioastronomie in Bonn. Further observational details are given in Table 1. Useful data were obtained on the Westerbork-Effelsberg baseline in both sessions, on the Chilbolton-Effelsberg baseline for 2 h in January (the source was not detected during the rest of the observation) and on the Knockin (Jodrell Bank) - Effelsberg baseline for the full track in June. Loci of projected baselines for the various interferometers are shown in Fig. 1.

Measurements of calibrator sources were made at regular intervals throughout network periods encompassing the SS433 observations. Based on the flux densities measured at Westerbork and Effelsberg, these data were used to calibrate the performance of the individual interferometers³. In the June session, the unresolved source 2134 + 004 (ref. 4) at declination 0.4° was observed at several intervals during the first and last 3 h of its apparition, to calibrate directly the changing gain, aperture blockage and spillover effects of the various antennas as a function of elevation for SS433 (at dec 4.9°). It was assumed that any differences in the general atmospheric conditions were negligible between observations of SS433 and 2134 + 004. The resulting calibration factors had an r.m.s. scatter of $\sim 5\%$. The point-to-point amplitude scatter for 2 min coherent integrations on the calibrator sources was $\sim 3\%$.

The flux densities of SS433 at 6 cm at the epochs of our measurements have been interpolated from measurements at 11 and 3.7 cm carried out by Johnston *et al.* (personal communication) assuming that the spectrum is a power law⁵. At these epochs, SS433 is classified as being in an inactive phase by Johnston *et al.*, meaning that the radio flux was changing by $< 20\%$ over a period of days. There is evidence that on 1–2 June, the flux density may have decreased by 8% during the period of our observations. We see little evidence of this in our visibility data, and it seems likely that variations of the visibility amplitude with hour angle are due to changes in the projected source structure rather than variations in the flux density.

The calibrated visibility amplitudes for the two sessions are shown in Fig. 2. Most of the information on source structure comes from the Westerbork-Effelsberg baseline, but the longer baselines provide important constraints on the structure. Although there are insufficient data for hybrid maps to be made, the relatively simple fringe amplitude curves allow several basic conclusions to be drawn about the structure and its variation. These conclusions are supported by the results of fitting simple models consisting of one or two barely resolved components and one or two extended elliptical gaussian components to the available data at each of the two epochs. Table 2 gives the parameters of the models and the range of values of each parameter which are consistent with the data. Curves representing the visibility amplitudes of these models are also shown in Fig. 2, and pictorial representations of the models from our data are given in Fig. 3.

The following deductions about the structure can be made.

(1) The size of the amplitude variations on these baselines

Table 1 Some parameters of the observations

Date	UT	Telescopes	Frequency (MHz)	Flux density (mJy)
9 January 1980	0600–1500	Westerbork Effelsberg Chilbolton	4,997	500 ± 20
1–2 June 1980	2045–0700	Westerbork Effelsberg Knockin	5,008.8	330 ± 30

The polarization on the sky was RCP in January and LCP in June.

shows the existence of structure on the 100–200 m arc s scale in both January and June.

(2) In January, the source was completely resolved on the longer projected spacings of the Chilbolton–Effelsberg baseline. The data shown for this baseline are for spacings similar to the Westerbork to Effelsberg baseline at extreme hour angles (see Fig. 1). In June, a barely resolved component of ~100 mJy is evident on the Knockin–Effelsberg baseline and on the Westerbork–Effelsberg baseline for the last 5 h.

(3) The projected baseline at which the peak fringe amplitude occurred for the Westerbork–Effelsberg interferometer differed in position angle (PA) by ~16° between the two observations. This indicates that the dominant PA of the 100–200 m arc s structure changed from $98^\circ \pm 3^\circ$ on 9 January (when it was aligned with the bulges in W50) to $114^\circ \pm 3^\circ$ on 1–2 June.

The following four points refer to the Westerbork–Effelsberg interferometer.

(4) An inflexion near IHA-6.5 h in the June curve, together with the best fitting model listed in Table 2, suggests that although the major part of the June emission was aligned in PA 114°, there is weaker emission still present in the 'fundamental' 98° PA.

(5) The width of the peaks in both curves, as well as their asymmetry, indicate that the structure is non-linear and very elongated.

(6) The difference in shape between the two curves indicates that the general appearance of the radio source changed appreciably between the two observations.

(7) The amplitudes of the maxima indicate that at both epochs the flux density from the compact (≤ 150 m arc s) structure was ~100 mJy less than the total flux density measured for SS433. This 'missing' 100 mJy can reasonably be ascribed to the larger (~3 arc s) structure known to exist from other measurements^{6,7}.

The change in orientation of the compact radio structure may be compared with the predictions of models for SS433 derived from optical observations. On the basis of Abell and Margon's⁸ model, the PA of the precessing jets seen in projection on the sky is expected to vary approximately sinusoidally with time, with a peak-to-peak amplitude of 41.4° and with the same period as the optical emission lines if the jets describe a cone of half angle 20.4° about a mean inclination of 79.7° to the line of sight^{9,10}. Alternatively, the PA has a saw-tooth variation with time (due to the 180° ambiguity in PA) with peak-to-peak amplitude of 172.8° if the angles are interchanged.

The position angles found from our data, and that of Walker *et al.*² on 12 March, 1979 (92° , range 85° – 105°) were compared with the expected behaviour on Abell and Margon's model and found consistent with precession with the smaller cone angle, as predicted by Begelman *et al.*¹¹. If the axis of precession projects onto the sky in position angle 98° (corresponding to the bulges in W50^{10,12,13}) then the PA of radio emission on the scale of 100 m arc s would lag the optical position angle by 5 ± 4 days, and would be phased so that the PA was decreasing in June. Although probably fortuitous, this agrees with the expected delay of about 4.5 days for the radio emitting clouds to reach an angular distance of 50 m arc s, assuming a velocity of 0.26c (ref. 4) and a distance^{10,12,14} of SS433 of 4 kpc (50 m arc s corresponds to a length of 6×10^{15} cm at a distance of 4 kpc).

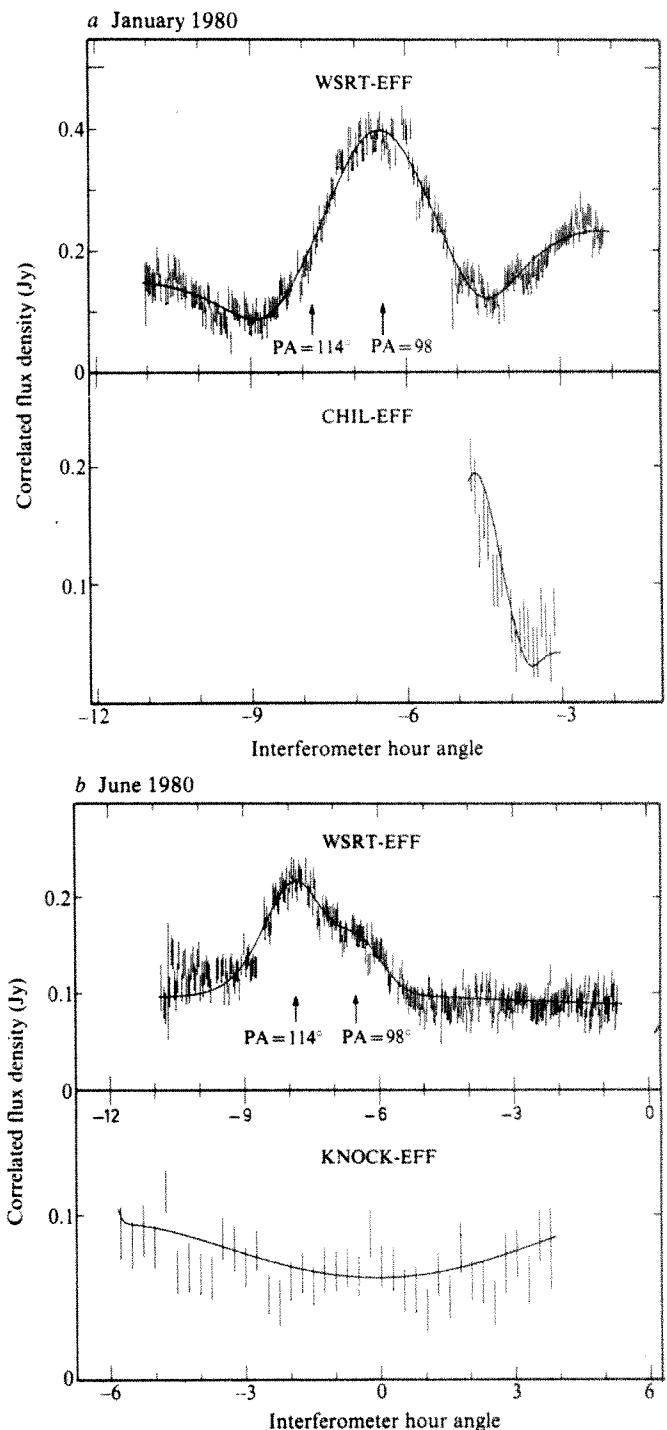


Fig. 2 The visibility amplitudes for SS433 at 6 cm, as a function of interferometer hour angle. *a*, January 1980, $S_T = 0.5$ Jy; *b*, June 1980, $S_T = 0.33$ Jy. Arrows indicate the hour angles at which a maximum in visibility amplitude is expected for PA 98° and 114° . WSRT, Westerbork; EFF, Effelsberg; CHIL, Chilbolton; KNOCK, Knockin. Errors bars are 2σ .

Our observations provide some evidence for extension of the radio source over a range in PA of $\pm 5^\circ$ to $\pm 10^\circ$ about the dominant PA. The presence of multiple-peaked spectral features have been noted^{9,15} in optical data which could be evidence of remnant emission from previous positions of the beam during its precessional motion. It is not known whether the nonlinearity in the compact radio structure is also due to remnant emission and whether it corresponds to the possible optical remnant emission. Any correspondence will depend on the mechanism for ejecting material into the beams, and also on the lifetimes of the radiating electrons in the different wavelength domains. These factors will affect the persistence of the emission. In the radio domain, the non-thermal integrated spectrum and the high

Table 2 Parameters of the models

Component	Flux (mJy)	Radial distance (m arc s)	PA (deg)	Angular size (m arc s)	Component PA (deg)
<i>a</i> January 1980					
1	89 ± 25	0	0	(92 ± 12) × (1 ± 1)	101 ± 1
2	219 ± 1	10 ± 1	99 ± 3	(22 ± 1) × (1 ± 1)	81 ± 2
3	101 ± 3	65 ± 1	99 ± 1	(22 ± 1) × (1 ± 1)	89 ± 4
<i>b</i> June 1980					
1	134 ± 5	0	0	(163 ± 5) × (12 ± 2)	119 ± 1
2	99 ± 1	0	0	(6.6 ± 0.5) × (5 ± 0.5)	98 ± 1
3	56 ± 5	0	0	(164 ± 5) × (4 ± 1)	98 ± 1

The columns give the component number, the flux density, the radial distance from component 1, the position angle of the location of the component (N to E) with respect to component 1, the major and minor axis, and the position angle of the component. The errors given are formal errors from the model-fitting process; the true errors are probably larger.

brightness temperatures of the compact components, ranging up to 10^9 K (ref. 2), imply that the emission is synchrotron radiation from relativistic electrons. The optical emission line system from the beams seems to arise from gas at a temperature of $<10^4$ K (ref. 11). There are arguments¹⁵ that the lifetime of the optical line emission is ≈ 10 days, but in the radio case there is not enough structural information on which to base calculations of synchrotron lifetimes. These relative lifetimes will also have some bearing on the lag of the rotation of the main axis of the radio emission behind that of the optical.

Although the optical data predict that during the January and June observations, the beams of SS433 lie within 10° of the plane of the sky, the detailed radio structures are different. Further data are required before patterns (if any exist) can be found in the variation of the structure. Radio flares will affect the structure for some time afterwards; note that the January data were taken 8 days after flaring activity in SS433 has ceased while the June data were taken 11 days before the end of the same long

period of quiescent behaviour (K. J. Johnston, personal communication). Perhaps the enhanced compact component in the June data can be identified with a pre-flare buildup of relativistic particles.

We have shown that the 100–200 m arc s structure in SS433 varies in orientation. Although the behaviour is consistent with the precessing beams model, further observations are needed to establish that the change in position angle is indeed periodic. Future comparison of variations in the radio structure and in the optical line emission may well provide unique information about the SS433 phenomenon.

We thank the staffs of the various observatories for their assistance, in particular R. S. Booth and P. N. Wilkinson at the UK telescopes; R. W. Porcas, H. Blaschke and U. Stursberg for advice and assistance during the processing of the data; and K. Johnston for flux density information. The Westerbork Synthesis Radio Telescope is operated by the Netherlands Foundation for Radio Astronomy with the financial support of the Netherlands Organization for Pure Research (ZWO).

Note added in proof: Hjellming and Johnston¹⁶ report extensive VLA observations of SS433 at seven epochs from 8 September 1979 to 20 June 1980. Their maps show an unresolved (≈ 100 m arc s) core source plus elongated 'jet' components up to several arc s in extent. Several of their results bear on the present paper. (1) The PA of the arc s scale structure remains within $\pm 20^\circ$ of 100° , which we have also deduced from our data with reference to the Abell and Margon model. (2) Proper motions observed for several features in the VLA maps have been extrapolated back to show that the radio jets are probably associated with the optical jets in the inner regions. There is some evidence from their data that the radio precession ephemeris lags that of the optical by 10–20 days with an error of the same magnitude. This is consistent with our result. (3) Assuming the core source is spherical, equipartition arguments lead Hjellming and Johnston to synchrotron loss lifetimes of 10^9 s, from which it is concluded that these losses cannot be significant in the core. The structure we derive for the central source is highly elongated implying higher magnetic fields and shorter synchrotron lifetimes. We do not believe that the structural information on the small scales is yet sufficiently good to confidently rule out one or another loss mechanism.

Received 15 December 1980; accepted 29 January 1981.

- Schilizzi, R. T., Norman, C. A., van Breugel, W. & Hummel, E. *Astr. Astrophys.* **79**, L26–L27 (1979).
- Walker, R. C. *et al. Astrophys. J.* (in the press).
- Cohen, M. H. *et al. Astrophys. J.* **201**, 249–255 (1975).
- Pauliny-Toth, I. I. K. *et al. Astr. J.* (submitted).
- Sequist, E. R., Gilmore, W., Nelson, G. J., Payten, W. J. & Slee, O. B. *Astrophys. J. Lett.* **241**, L77–L81 (1980).
- Spencer, R. E. *Nature* **282**, 483–484 (1979).
- Gilmore, W. & Sequist, E. R. *Astr. J.* **85**, 1486–1495 (1980).
- Abell, G. O. & Margon, B. *Nature* **279**, 701–703 (1979).
- Margon, B., Grandi, S. A. & Downes, R. A. *Astrophys. J.* **241**, 306–315 (1980).
- Geldzahler, B. J. thesis, Univ. Pennsylvania (1980).
- Begelman, M. C., Sarazin, C. L., Hatchett, S. P., McKee, C. F. & Arons, J. *Astrophys. J.* **238**, 722–730 (1980).
- Geldzahler, B. J., Pauls, T. & Salter, C. J. *Astr. Astrophys.* **84**, 237–244 (1980).
- Seward, F., Grindlay, J., Sequist, E. & Gilmore, W. *Nature* **287**, 806–808 (1980).
- Van Gorkom, J., Goss, W. M. & Shaver, P. A. *Astr. Astrophys.* **82**, L1–L2 (1980).
- Murdin, P., Clark, D. H. & Martin, P. G. *Mon. Not. R. astr. Soc.* **193**, 135–151 (1980).
- Hjellming, R. M. & Johnston, K. J. *Nature* **290**, 100–107 (1981).

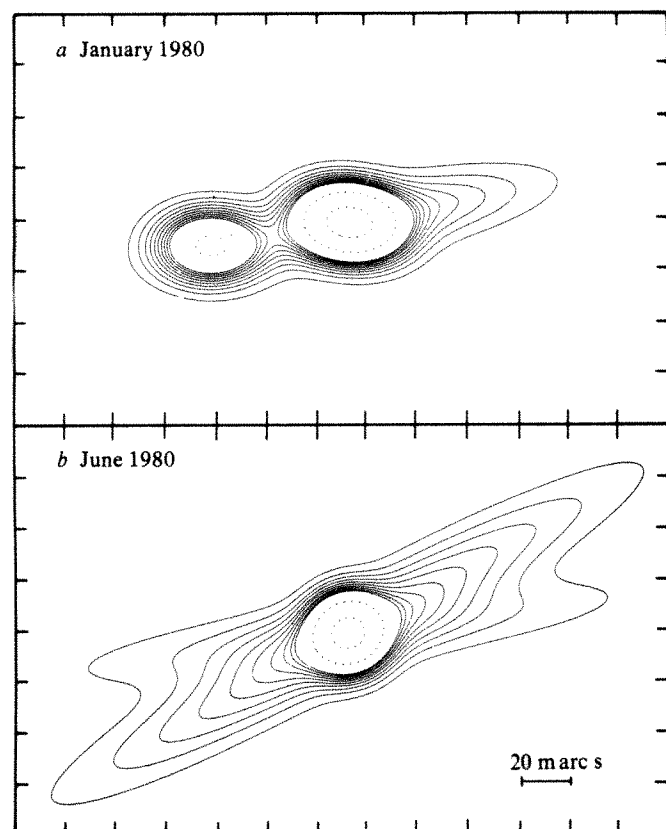


Fig. 3 Models for SS433 at 6 cm. *a*, 9 January 1980; *b*, 1–2 June 1980. There is a 180° ambiguity in the orientation of these diagrams because they were derived using amplitude data alone. North or south is up. The data have been smoothed with a 20 m arc s circular gaussian beam. The contour levels are: January 1.5, 3, 4, 5, ... 15% (solid), 45, 90% (dashed); June, 2, 4, 6, ... 20% (solid), 40, 80% (dashed).

Acetonitrile in the stratosphere—implications from laboratory studies

H. Böhringer & F. Arnold

Max-Planck-Institut für Kernphysik, PO Box 103980, D6900 Heidelberg, FRG

The presence of acetonitrile (CH_3CN) in the stratosphere has been tentatively inferred¹ from measurements of stratospheric positive ions around 36 km. As well as the expected proton hydrates $\text{H}^+(\text{H}_2\text{O})_n$, another ion family was observed to which was assigned the general form, $\text{H}^+\text{X}_l(\text{H}_2\text{O})_m$, where X is a molecule of mass 41 ± 1 AMU. If these ions are formed from proton hydrates by reactions involving the unknown trace gas X, the latter must have a proton affinity of ≥ 175 kcal mol⁻¹ and a number density of the order of 10^5 cm⁻³ around an altitude of 36 km. Two previously proposed identifications for X—NaOH and MgOH—have recently been excluded by new stratospheric ion composition data^{3,4} which unambiguously determined the mass of X to be 41 AMU and failed to show $\text{H}^+\text{X}_l(\text{H}_2\text{O})_m$ ions containing the Mg isotopes 25 and 26. Another suggestion¹ that X might be acetonitrile (CH_3CN) has been criticized on the grounds that it is difficult to explain its presence in the stratosphere². We now report laboratory studies of clustering equilibria for $\text{H}^+(\text{CH}_3\text{CN})_l(\text{H}_2\text{O})_m$ ions, which further support identification of X as acetonitrile.

The measurements were made using a newly designed temperature-variable drift tube mass spectrometer which will be described in detail elsewhere⁵. In this apparatus the primary ions are produced in a separate ion source and are mass selected before they are introduced into the drift tube. The ions traverse the tube drifting in an electric field while they are subjected to reactions with reactant gases. Having passed the drift tube exit aperture the product ions are mass analysed. During the measurements described here the drift tube was operated at low electric fields (≤ 2 V cm⁻¹) so that the additional 'drift energy' of the ions was $< 2\%$ of the thermal energy and only minor temperature corrections have been applied for this effect. The gas pressures applied to the drift tube ranged from 0.2 to 3 torr and the temperatures from 220 to 400 K. Hydrogen was used as buffer gas to which the reactant gas CH_3CN or H_2O was added in amounts of 0.2–10% and the second of the two gases was only added in traces if mixed ligand clusters were to be observed. Primary H_2^+ ions were injected into the drift tube and rapidly converted to H_3^+ by reactions with hydrogen. Proton transfer reactions of these ions led to the core ions H_3O^+ and CH_3CN^+ , which were subjected to further clustering and ligand exchange reactions. We assumed that chemical equilibrium had been established when a further increase of the buffer gas pressure and the reaction time did not change the ratio of product to precursor ion abundance. Absorption and desorption of the reactant gas at the walls of the drift tube made control of the reactant gas pressure difficult and introduced an estimated error of up to a factor of 3 in the measured equilibrium constant.

The $\text{H}^+(\text{CH}_3\text{CN})_l(\text{H}_2\text{O})_m$ ions produced in the laboratory showed all the ion masses observed in the stratosphere by Arnold *et al.*^{1,3}, and essentially all ion masses detected by Arijs *et al.*^{4,6} at an altitude around 34 km (see Table 1). The measured equilibrium constants for association of H_2O and CH_3CN to different cluster ions are given in Fig. 1 for a temperature of 227 K (the ambient temperature during the stratospheric measurement in ref. 3). A comparison of the equilibrium constants for the proton hydrates with previous high-pressure ion source measurements⁷ also given in Fig. 1 shows a fairly good agreement.

The results show that hydration equilibrium constants for heteromolecular clusters containing CH_3CN molecules are generally lower than those for $\text{H}^+(\text{H}_2\text{O})_n$ ions and they decrease with increasing number of CH_3CN ligands contained in the

complex. The equilibrium constants for the $\text{H}^+(\text{CH}_3\text{CN})_p$ ions with $p \geq 3$ are remarkably low despite the large dipole moment of CH_3CN . Thus, it seems that the ligand interaction between H_2O and CH_3CN and among CH_3CN molecules is weaker than that between H_2O molecules which is dominated by hydrogen bonding. This is consistent with the lower heat of evaporation of CH_3CN compared with H_2O .

A striking exception, however, is found for the ion $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$ which is the most stable ion containing four ligands and has a much larger equilibrium constant than the two precursors, $\text{H}^+(\text{CH}_3\text{CN})_3$ and $\text{H}^+(\text{CH}_3\text{CN})_2\text{H}_2\text{O}$. Using collision-induced dissociation, we found that decomposition of $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$ preferentially leads to the fragment $\text{H}^+(\text{CH}_3\text{CN})_2\text{H}_2\text{O}$, implying that the water molecule is bound more strongly than the acetonitrile molecules. Therefore, the high stability of this ion may be explained by a structural rearrangement, in which H_3O^+ forms the core ion and the acetonitrile molecules are symmetrically attached to the three hydrogen atoms sharing the positive charge. A similarly enhanced stability was found for the ion $\text{H}^+(\text{CH}_3\text{CN})_4(\text{H}_2\text{O})_2$ among the ions with six ligands. An analogous symmetric structure can be attributed to this ion in which the two water molecules attach to the proton leaving four, partly charged H-atom bonding sites for the acetonitrile molecules.

In the stratosphere, hydration is sufficiently fast to attain hydration equilibria for all positive ions. Therefore the hydration equilibria of stratospheric X-containing and laboratory CH_3CN -containing ions can be compared—as shown in Table 2—and a fairly good agreement is found. The laboratory results indicate that among the $\text{H}^+(\text{CH}_3\text{CN})_3(\text{H}_2\text{O})_k$ family, $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$ should be the only detectable atmospheric ion, which is, in fact, shown by the stratospheric observations.

The stratospheric observations of H_2O -clustering equilibria are therefore consistent with X being acetonitrile. The only

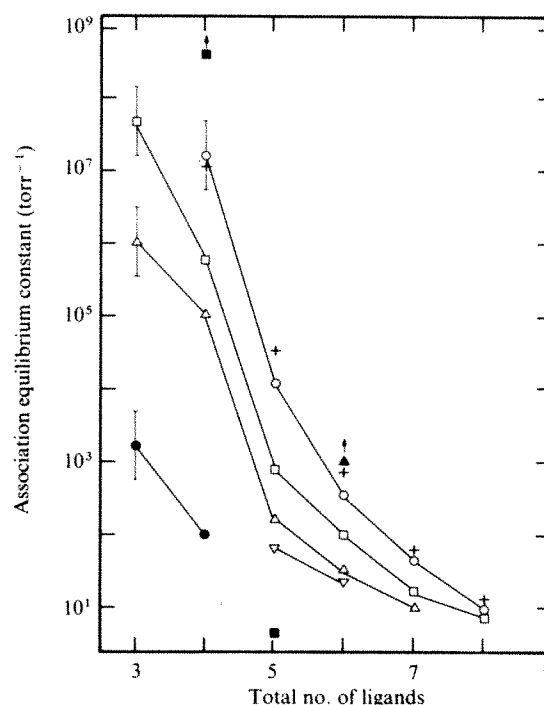


Fig. 1 Measured equilibrium constants for H_2O and CH_3CN association reactions at 227 K. The equilibrium constants $\geq 10^4$ (torr⁻¹) had to be extrapolated from higher temperatures. For $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$ and $\text{H}^+(\text{CH}_3\text{CN})_4(\text{H}_2\text{O})_2$ only lower limits can be given. The results for the proton hydrates are compared with measurements of Kebarle *et al.*⁶. The estimated error of up to a factor of 3 valid for all measured values is indicated by error bars for the first points of the curves. $\circ$, $\text{H}^+(\text{H}_2\text{O})_n$; $\square$, $\text{H}^+\text{CH}_3\text{CN}(\text{H}_2\text{O})_m$; $\triangle$, $\text{H}^+(\text{CH}_3\text{CN})_2(\text{H}_2\text{O})_l$; ∇ , $\text{H}^+(\text{CH}_3\text{CN})_3(\text{H}_2\text{O})_k$; $\bullet$, $\text{H}^+(\text{CH}_3\text{CN})_4(\text{H}_2\text{O})_2$; $\blacksquare$, $\text{H}^+(\text{CH}_3\text{CN})_3$; $\blacktriangle$, $\text{H}^+(\text{CH}_3\text{CN})_4$; $+$, $\text{H}^+(\text{H}_2\text{O})_2(\text{CH}_3\text{CN})_x$ (ref. 6).

exception is the observed ion with mass 136 ± 1 AMU. Assuming that the previous identification, $\text{H}^+(\text{CH}_3\text{CN})_2(\text{H}_2\text{O})_3$, is correct, the ion should be one or two orders of magnitude less abundant than observed. Therefore, this ion could result from reactions involving another trace gas, Y, and thus have the form $\text{H}^+\text{Y}(\text{H}_2\text{O})_i$ or $\text{H}^+\text{YX}(\text{H}_2\text{O})_j$. If this cluster ion contains three or four ligands like the other observed heteromolecular clusters, the mass of the molecule Y should be 99 ± 1 , 81 ± 1 , 76 ± 1 or 58 ± 1 AMU. Possible candidates for Y are CH_3NO_3 (mass 77), $\text{C}_2\text{H}_5\text{CHO}$ (mass 58), and $(\text{CH}_3)_2\text{CO}$ (mass 58) which all have proton affinities in excess of $175 \text{ kcal mol}^{-1}$ and therefore should replace water molecules in proton hydrates in stratospheric conditions. Note that a recently detected⁴ ion of minor abundance with mass 117 AMU may belong to the same hydrate ion family as the ion 136 ± 1 AMU. If so, the mass of Y should be 99, 80, 75 or 57 which would exclude all identifications suggested above.

Table 1 Observed stratospheric positive ions around 36 km altitude

Mass	Identification	Ion mass for $\text{X} = \text{CH}_3\text{CN}$	Measured abundance (%)	
			Ref. 3	Ref. 4
55	$\text{H}^+(\text{H}_2\text{O})_3$	55	3.4	5.5
73	$\text{H}^+(\text{H}_2\text{O})_4$	73	26.8(21.3)*	29.6
78	$\text{H}^+\text{X}(\text{H}_2\text{O})_2$	78	7.2	8.4
91	$\text{H}^+(\text{H}_2\text{O})_5$	91	8.9(17.8)*	2.8
96	$\text{H}^+\text{X}(\text{H}_2\text{O})_3$	96	19.6	44.4
101	$\text{H}^+\text{X}_2\text{H}_2\text{O}$	101	10.7	3.5
114	$\text{H}^+\text{X}(\text{H}_2\text{O})_4$	114	—	0.9
119 ± 1	$\text{H}^+\text{X}_2(\text{H}_2\text{O})_2$	119	16.1	4.9
136 ± 1	$\text{H}^+\text{X}_2(\text{H}_2\text{O})_3$	137	4.4	—
141 ± 1	$\text{H}^+\text{X}_3\text{H}_2\text{O}$	142	2.9	—

* Values in parentheses are corrected for dissociation (see ref. 3).

Finally, implications of the measured CH_3CN clustering equilibria will be considered. Arnold *et al.*³ have applied a steady-state treatment of the reactions converting the proton hydrates to cluster ions, containing 1, 2 and 3 X molecules, to their stratospheric ion composition data. They found that the ion family containing three X ligands is one order of magnitude less abundant than expected, assuming equal reaction rates for all three conversion steps. This was tentatively explained by a process converting the only observed ion with three CH_3CN molecules, $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$, back to cluster ions containing two CH_3CN ligands. The process should be about 10 times faster than ion-ion recombination in the stratosphere around 36 km. The laboratory studies revealed that the CH_3CN molecules are less strongly bound than H_2O in the $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$ complex. Therefore, two processes could be responsible for the backward conversion: thermal decomposition or ligand exchange of CH_3CN for a H_2O molecule. For the thermal dissociation, a free enthalpy value of $\Delta G \approx 13.8 \text{ kcal mol}^{-1}$ is calculated for 227 K and a gas density of $4 \times 10^{16} \text{ cm}^{-3}$ to account for the observed ion abundances, while for the ligand exchange reaction $\Delta G \approx 5.8 \text{ kcal mol}^{-1}$ is obtained for the concentration ratio $[\text{H}_2\text{O}]/[\text{X}] = 3 \times 10^6$ (ref. 3). The corresponding values derived from the laboratory measurements for the CH_3CN -containing clusters are $\Delta G \geq 12 \text{ kcal mol}^{-1}$ and $\Delta G \geq 3.6 \text{ kcal mol}^{-1}$. A thermal dissociation or ligand exchange reaction affecting the $\text{H}^+(\text{CH}_3\text{CN})_3\text{H}_2\text{O}$ ion is also observed in the laboratory studies at large $[\text{H}_2\text{O}]/[\text{CH}_3\text{CN}]$ ratios ($\gg 10^3$). We therefore conclude that the ΔG values for both processes hardly exceed the lower limits given above by a large factor, which agrees well with the findings from the atmospheric observations.

Thus the observed masses and abundances of positive stratospheric ions can be well explained by proton hydrates and $\text{H}^+(\text{CH}_3\text{CN})_i(\text{H}_2\text{O})_m$ ion clusters. The ions 136 ± 1 and 117 AMU which do not fit into this scheme might reflect the influence of a stratospheric trace gas other than acetonitrile.

Until now, no source of stratospheric acetonitrile has been known. The sharp decrease in fractional abundances of X-

Table 2 Relative comparison of hydration equilibria of laboratory and stratospheric ions

Abundance ratio	Laboratory data	Ref. 3	Ref. 4
[96]/[78]	$1.7 \times 10^{-11} [\text{H}_2\text{O}]$	2.7	5.3
[119]/[101]	$3.5 \times 10^{-12} [\text{H}_2\text{O}]$	1.5	1.4
[91]/[73]	$3.5 \times 10^{-13} [\text{H}_2\text{O}]$	0.84	0.1
[114]/[96]	$2.2 \times 10^{-14} [\text{H}_2\text{O}]$	—	0.02
[136 ± 1]/[119]	$4.7 \times 10^{-15} [\text{H}_2\text{O}]$	0.27	—

Ions are referred to by their mass number.

containing ions above 40 km observed in rocket experiments⁸ indicates that the source of X should be at heights below about 35 km. Further ion composition measurements covering a larger height range may help locate this source and provide additional thermodynamic data on X-containing ions which can be compared with our laboratory results.

We thank Professor E. E. Ferguson for helpful discussions. Part of the project was funded by the Bundesminister für Forschung und Technologie through GSF.

Note added in proof: The kinetics of the switching reactions of proton hydrates with CH_3CN has been studied by Smith *et al.*⁹

Received 1 December 1980; accepted 12 February 1981.

1. Arnold, F., Böhringer, H. & Henschen, G. *Geophys. Res. Lett.* **5**, 653–656 (1978).
2. Ferguson, E. E. *Geophys. Res. Lett.* **5**, 1035–1038 (1978).
3. Arnold, F., Henschen, G. & Ferguson, E. E. *Planet. Space Sci.* **29**, 185–193 (1981).
4. Arijis, E., Nevejans, D. & Ingels, J. *Nature* **288**, 684–686 (1980).
5. Böhringer, H. & Arnold, F. *Int. J. Mass Spectrom. Ion Phys.* (submitted).
6. Arijis, E., Ingels, J. & Nevejans, D. *Nature* **271**, 642–644 (1978).
7. Kebarle, P., Searles, S. K., Zolla, A., Scarborough, J. & Arshadi, M. *J. Am. chem. Soc.* **89**, 6393–6399 (1967).
8. Arnold, F., Krankowsky, D. & Marien, K. H. *Nature* **267**, 30–32 (1977).
9. Smith, D., Adams, N. G. & Alge, E. *Planet. Space Sci.* (in the press).

Unusual *PT* dependence of thermal conductivity for a clathrate hydrate

R. G. Ross, P. Andersson & G. Bäckström

Department of Physics, Umeå University, S-901 87 Umeå, Sweden

Molecular clathrate solids¹ are crystals with peculiar and interesting properties. Their characteristic feature² is that one chemical species, the 'host', forms a strongly bonded structure, which encages individual molecules of a second 'guest' species. There is virtually no chemical bonding between host and guest molecules. The guests are, nevertheless, firmly constrained by high energy barriers. X-ray and neutron diffraction experiments have shown that clathrates have a periodic structure typical of crystalline solids. A clathrate hydrate² is a host substance consisting of H_2O molecules which form an ice-like, hydrogen-bonded structure. Huge geological deposits of clathrate hydrates containing natural gas are known to exist^{2,3}—these deposits may be exploitable, and in searching for clathrate hydrates knowledge of the thermal conductivity, λ , will be important. We present here a preliminary account of the first measurements of λ versus pressure (*P*) and temperature (*T*) for a clathrate hydrate. We studied a clathrate hydrate encaging molecules of tetrahydrofuran. Thermal conductivity was measured using the transient hot-wire method⁴. The accuracy of the measurements was 3% and the precision 1%. The temperature dependence was found to be unusual in that $\lambda(T)$ has a positive slope (Fig. 1), a feature never before observed with crystalline organic materials. Furthermore, when λ was measured versus *P*, it was found unexpectedly that $\lambda(P)$ was constant—behaviour never before reported for a non-metallic material.

The temperature dependence almost universally observed at $T > 100 \text{ K}$ with crystalline, non-metallic solids is $\lambda(T) \propto T^{-1}$. This negative slope agrees with the theory of phonon-phonon scattering by the simplest umklapp mechanism. Positive slopes

of $\lambda(T)$ are encountered for non-crystalline materials in the glass state. Here topological structural disorder is considered to provide the dominant phonon scattering mechanism. Disorder fixes the phonon mean free path, and the positive slope of $\lambda(T)$ is attributed to variation of the heat capacity in the well-known formula for λ due to Debye. A few inorganic crystalline solids have positive slopes of $\lambda(T)$, that is YB_{68} , KFeF_3 and FeAl_2O_4 , and structural disorder has been invoked to account for this behaviour⁵⁻⁷. Orientationally disordered plastic crystal phases⁸ are intermediate cases, showing a negative slope of $\lambda(T)$ but with much reduced steepness: $\lambda \propto T^{-0.2}$.

All non-metallic materials previously investigated exhibit a finite slope of $\lambda(P)$, usually positive, as would be expected. A few cases of negative slope have, however, recently been discovered, one being ordinary ice⁹. The magnitudes of the slopes are most conveniently compared in terms of $g = \partial \ln \lambda / \partial \ln \rho$, where ρ is the mass density. For most cases g ranges from 3 to 10, whereas for exceptional materials $g = -2.6$ (ice Ih)⁹ and $g = -6.2$ (NH_4F)^{10,11}. From the compressibility of a propane clathrate hydrate^{2,3}, on the other hand, we estimate $|g| < 0.2$, which is an order of magnitude smaller than for other substances.

The thermal conductivity of the clathrate was found to be only 15% greater than for the corresponding liquid solution. This is in qualitative agreement with the only previous measurement for a propane hydrate³. The increase in λ when going from liquid to an orientationally disordered plastic crystal phase⁸ is typically 5%, whereas a dramatic increase is usually recorded for a normal crystal¹².

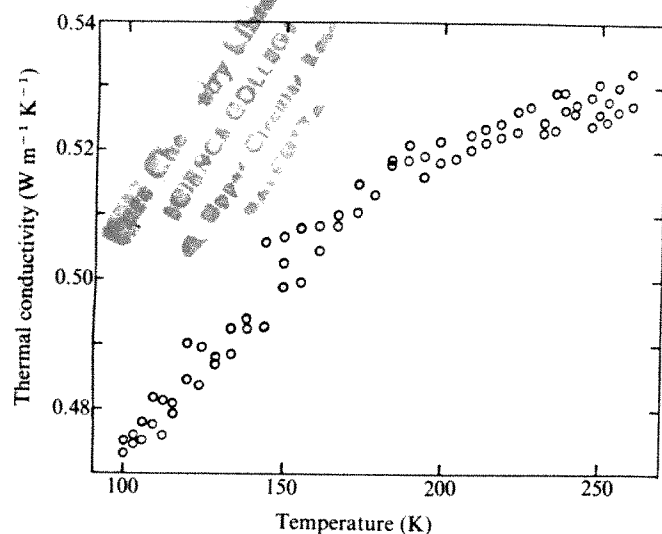


Fig. 1 Isobaric variation with temperature of thermal conductivity, λ , of (tetrahydrofuran) · 17.05(H_2O) clathrate hydrate at a pressure of 0.1 GPa.

Although the tetrahydrofuran clathrate hydrate is crystalline, there is in fact some structural disorder². Recent neutron inelastic scattering results¹³ indicate reorientational motion of the guest molecules with an activation energy of ~ 0.01 eV. At temperatures of 100–300 K there would thus be frequent reorientation of the guest molecules, which might provide a phonon scattering mechanism. We propose that this dynamical structural disorder causes a dominant thermal resistivity in our clathrate hydrate. From the pressure variation of λ we conclude that the total resistive mechanism is virtually insensitive to changes in density, but we cannot yet exclude that this constancy is a combined effect of positive and negative contributions from the main scattering mechanisms.

The phase diagram for mixtures of tetrahydrofuran and H_2O versus P , T and composition is well known¹⁴. The so-called structure II clathrate hydrate exists in a very narrow region of composition. At atmospheric pressure it occurs at a composition of (tetrahydrofuran) · 17(H_2O) with congruent melting. We prepared our specimens by pouring a liquid solution of double-

distilled H_2O and spectroscopic-grade tetrahydrofuran into the Teflon measurement cell of a piston-cylinder high pressure apparatus. On decreasing the temperature at 0.01 GPa freezing occurred after 5 K of supercooling and was detected by a marked thermal arrest at 277.0 K in agreement with the phase diagram¹⁴. Such thermal behaviour during solidification excludes the possibility of glass formation. The observed discontinuity of λ at the arrest temperature corroborates the conclusion that a first-order transition was induced on cooling. Experiments were carried out at four different compositions from 16.59 to 17.55 moles of H_2O , and the variation of λ versus T and P remained the same, although the absolute values of λ differed by as much as 20% between the extreme cases. It is thus unlikely that the variation of λ reported here could have been caused by phase separation in response to changes in P and T .

It is reasonable to assume that natural-gas clathrate hydrates would behave similarly to the substance investigated here. One could thus expect λ to be small and to change little over geologically relevant ranges of temperature and pressure.

We thank Dr D. W. Davidson for suggesting this investigation and Dr G. A. Slack for useful discussion. The work was financially supported by the Swedish Natural Science Council.

Received 8 December 1980; accepted 27 January 1981.

1. Parsonage, N. G., & Staveley, L. A. K. *Disorder in Crystals* (Clarendon, Oxford, 1978).
2. Davidson, D. W. in *Water: A Comprehensive Treatise* Vol. 2 (ed. Franks, F.) 115–234 (Plenum, New York, 1973).
3. Stoll, R. D. & Bryan, G. M. *J. geophys. Res.* **84**, 1629–1634 (1979).
4. Ross, R. G., Andersson, P. & Bäckström, G. *Molec. Phys.* **38**, 377–385 (1979).
5. Slack, G. A., Oliver, D. W. & Horn, F. H. *Phys. Rev.* **B4**, 1714–1720 (1971).
6. Suemune, Y. *J. phys. Soc. Jap.* **19**, 2234 (1964).
7. Slack, G. A. *Phys. Rev.* **134**, A1268–A1280 (1964).
8. Andersson, P. & Ross, R. G. *Molec. Phys.* **39**, 1359–1368 (1980).
9. Andersson, P., Ross, R. G. & Bäckström, G. *J. phys. C: Solid State Phys.* **13**, L73–L76 (1980).
10. Ross, R. G. & Sandberg, O. *J. phys. C: Solid State Phys.* **11**, 667–672 (1978).
11. Swenson, C. A. & Tedeschi, J. R. *J. chem. phys.* **40**, 1141–1143 (1964).
12. Slack, G. A. *Solid State Phys.* **34**, 1–71 (1979).
13. Wegener, W., Vanderhaeghen, J., Hautecler, S., Legrand, E. & van Gervan, L. in *Neutron Inelastic Scattering 1977*, Vol. 1, 415–433 (IAEA, Vienna, 1978).
14. Dyadin, Yu. A., Kuznetsov, P. N., Yakovlev, I. I. & Pyrinova, A. V. *Doklady. Chem.* **208**, 9–12 (1973).

Destruction of polychlorodibenzo-*p*-dioxins

D. C. Ayres

Department of Chemistry, Westfield College, Hampstead, London NW3 7ST, UK

The release of 2,4,5-trichlorophenol (2,4,5-T) containing 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) at Sèvres¹ highlighted the need for an efficient control procedure for the highly toxic² dioxin. The concomitant formation of the dioxin and the trichlorophenol during hydrolysis of *sym*-tetrachlorobenzene has led to demands for restrictions on the use of 2,4,5-T and related herbicides³. An environmental hazard is also presented by pentachlorophenol⁴ which includes the more highly chlorinated dioxins, in particular octachlorodibenzo-*p*-dioxin (OCDD). The use of polychlorophenol formulations has been banned in Sweden on the grounds⁵ that heat-stable chlorodioxins are formed during incineration; this also points to the need for degradation by chemical means. Dehalogenation of polychlorodibenzodioxins (PCDD) can be accomplished by photolysis in laboratory conditions⁶, but material adsorbed on soil is little affected⁷. Cleavage of the ether linkages with the formation of halophenols may be achieved by treatment with strong acids or quaternary ammonium salts⁸. The dibenzodioxin nucleus, is however, rather resistant to chemical attack⁹ and in the absence of a suitable control procedure the use of 2,4,5-T has been restricted in the US by the Environmental Protection Agency^{10,11}. I now report the oxidative degradation of TCDD and related compounds by ruthenium tetroxide.

Ruthenium tetroxide is a powerful oxidizing agent which reacts with a wide range of organic substances¹² including those aromatic compounds which are oxidized slowly, or not at all, by permanganate¹³. It may be safely used in solution in water or in organic solvents with no nucleophilic character, such as chloroform, nitromethane or carbon tetrachloride.

In a preliminary experiment we showed that 2,7-dichlorobenzo-*p*-dioxin (DCDD) is destroyed by ruthenium tetroxide in carbon tetrachloride solution on standing overnight at ambient temperature. The progress of this and related oxidations may be monitored by thin layer or preferably by gas chromatography, where a 2-m column of 5% OV1 on Chromosorb 750 was operated at 195 °C, with a nitrogen flow of 50 ml min⁻¹ giving a retention time of 162 s.

Subsequent work on TCDD was carried out in collaboration with J. Harrison at the Chemical Defence Establishment, Porton Down. The progress of the oxidation of TCDD ($C = 72 \mu\text{g ml}^{-1}$) at ambient temperature in carbon tetrachloride containing the tetroxide ($2 \times 10^3 \mu\text{g ml}^{-1}$, about 10 equiv.) was monitored at intervals by stopping aliquots with benzene (2 ml). Residual concentrations were determined by comparison of their electron response with that of a standard solution of TCDD ($C = 60 \mu\text{g ml}^{-1}$) with the GLC operating at 230 °C. The half life of 560 min was evaluated from the linear pseudo-first order plot of

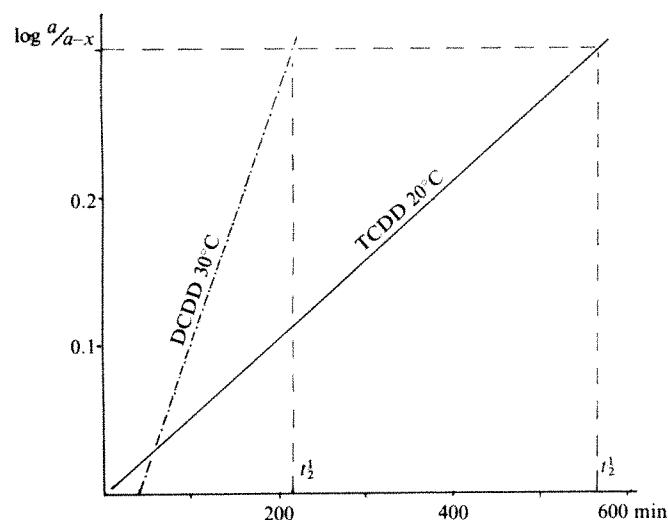


Fig. 1 Rates of oxidation of polychlorodibenzodioxins by ruthenium tetroxide.

$\log (\text{proportion residual})^{-1}$ to time (Fig. 1). The rate of oxidation was considerably increased at higher temperatures: in carbon tetrachloride at 70 °C and with the same concentration of reagents the half life of TCDD was <15 min.

For general guidance, within this group of compounds, the temperature coefficient for the oxidation of 2,7-DCDD was determined. The rate of oxidation of this non-toxic substance at ambient temperature is similar to that of TCDD; it can therefore act as a mimic for TCDD in degradative studies. The kinetic investigation was carried out in carbon tetrachloride solution containing DCDD ($70 \mu\text{g ml}^{-1}$) and ruthenium tetroxide ($4 \times 10^3 \mu\text{g ml}^{-1}$); the results given in Table 1 were obtained from the linear plots of $\log (\text{proportion residual})^{-1}$ to time as illustrated for TCDD.

The rate constants were obtained from the half lives after the induction period from the relation $k = 1/\ln 2 \cdot t_{1/2}$. The induction period diminished progressively as the reaction temperature rose and in these examples is probably associated with a free radical process, because the formation of radical-cations from dibenzo-*p*-dioxins has been demonstrated¹⁴ and these species have also been detected¹⁵ as intermediates in other oxidations or aromatic compounds by ruthenium tetroxide.

The values for the energy of activation were calculated from the three rate constants using the Arrhenius equation giving a mean value of about 68 kJ mol⁻¹. It follows from this that the rate increase per 10 °C rise in this temperature range is about 2.4-fold.

It is possible to show that all members of the PCDD group are susceptible to attack by ruthenium tetroxide. This follows because octachlorodibenzo-*p*-dioxin (OCDD), with the most inactivating chlorine substituents, is itself degraded in refluxing

Table 1 Temperature dependence of the oxidation of DCDD⁶ ruthenium tetroxide

Temperature (°C)	Induction period (min)	Total half life (min)	t_1 after induction (min)	k (min ⁻¹)	E_{ACT} (kJ mol ⁻¹)
30	37	215	178	3.89×10^{-3}	66.65
40	10	87	77	9.00×10^{-3}	69.1
50	4	38	34	2.04×10^{-2}	

carbon tetrachloride ($C = 70 \mu\text{g ml}^{-1}$) by the oxidant ($4 \times 10^3 \mu\text{g ml}^{-1}$), with a half life of ~10 min. This result is also significant because of the possible enzymatic reduction of OCDD to dibenzodioxins of lower chlorine number: analogous changes occur¹⁶ during the metabolism of hexachlorobenzene.

We have no evidence of the nature of the fragments formed during oxidation of the polychlorodibenzodioxins. However, the related chlorophenols undergo extensive decomposition to yield chloride ions and no significant levels of organic products¹⁷.

The present work shows that ruthenium tetroxide can be used both for the detoxification of glassware and artefacts, and for the periodical purging of industrial reactors to counteract the accumulation of polychloro-*p*-dioxin residues. When using the reagent note that there is some uncertainty in the total reaction time owing to the period of induction. The induction periods given in Table 1 were, however, reasonably reproducible and the delay is not significant at the higher temperature. To obtain information about reaction kinetics the oxidant was used in excess in carbon tetrachloride, but the method is also applicable to the oxidation of PCDDs with catalytic amounts of ruthenium tetroxide as the primary reagent and regeneration by a consumable secondary oxidant, such as hypochlorite.

We thank Dr J. Harrison for DCDD and TCDD and Dr A. J. Dobbs for OCDD. This paper is Part 6 in the series 'Oxidation of Aromatic Substrates'. It was presented in part at the Chemical Society Pollution Symposium in Belfast on 26 September 1979, and subsequently in London on 1 May 1980.

Received 28 October 1980; accepted 27 January 1981.

- Hay, A. *Nature* **281**, 521 (1979).
- Int. Agency Res. Cancer* **15**, 42 (1977).
- Crosland, J. *Ecologist* **10**, 87 (1980).
- Dickens, D. *Nature* **283**, 418 (1980).
- Jansson, B. & Sundström, G. *Sci. Total Envir.* **10**, 209 (1978).
- Dobbs, A. J. & Grant, C. *Nature* **278**, 163 (1979).
- Kearney, P. C., Woolson, E. A. & Ellington, C. P. *Envir. Sci. Technol.* **6**, 1017 (1972).
- Botré, C., Memoli, A. & Alhaque, F. *Envir. Sci. Technol.* **13**, 228 (1979).
- Blair, E. H. (ed.) *Chlorodioxins-Origin and Fate Adv. Chem. Ser.*, No. 120 (American Chemical Society, 1973).
- Cookson, C. *Nature* **278**, 108 (1979).
- Josephson, J. *Envir. Sci. Technol.* **14**, 1165 (1980).
- Lee, D. G. & van den Engh, M. in *Oxidation in Organic Chemistry* (ed. Trahanovsky, W. S.) (Academic, New York, 1973).
- Ayres, D. C. & Scott, C. M. *Envir. Sci. Technol.* **13**, 983 (1979).
- Shine, H. J. & Shade, L. R. *J. heterocycl. Chem.* **11**, 139 (1974).
- Ayres, D. C. & Gopalan, R. *JCS Chem. Commun.* 890 (1976).
- Mehendale, H. M., Fields, M. & Matthews, H. B. J. *Agric. Food. Chem.* **23**, 261 (1975).
- Ayres, D. C. & Levy, D. P. (in preparation).

A new pyroxene fractionation trend from a layered basic intrusion

J. R. Wilson*, K. H. Esbensen† & P. Thy‡

* Department of Geology, University of Aarhus, 8000 Århus C, Denmark

† Department of Metallurgy, Technical University of Denmark, Building 204, 2800 Lyngby, Denmark

‡ Nordic Volcanological Institute, University of Iceland, 101 Reykjavik, Iceland

Pyroxene fractionation trends in igneous rocks depend on the magma composition and the conditions of crystallization. Four main types of pyroxene crystallization trend have been recognized in terrestrial igneous rocks (Fig. 1). These are the clinopyroxene trend of mildly alkaline basic magmas^{1,2}; the clinopyroxene trend of strongly alkaline basic magmas³; the two-pyroxene trend typical of tholeiitic magmas⁴⁻⁶; and the two-pyroxene trend typical of calc-alkaline intrusions⁷⁻⁹. The calc-alkaline pyroxene trends show restricted iron-enrichment as high P_{H_2O} causes the crystallization of calcic amphibole (and biotite) in place of pyroxenes at a fairly early stage of fractionation. Ca-rich and Ca-poor pyroxenes from the Fongen-Hyllingen complex which coexist with hydrous mafic phases and initially have compositions very similar to those from calc-alkaline complexes, show extreme iron-enrichment in the late differentiates. This is the first recorded example of such extreme pyroxene fractionation in elevated P_{H_2O} conditions and this fractionation trend, accompanied by other distinct mineralogical features, warrants designation as a new type: the Fongen-Hyllingen trend.

The 160 km² Fongen-Hyllingen complex is an extremely differentiated basic intrusion, synorogenically emplaced in the Scandinavian Caledonides of the Trondheim region, Norway¹⁰, (11.5°E, 63°N). The complex was emplaced at a pressure of 5–6 kbar before the regional deformation and metamorphism in amphibolite facies¹¹ peaked. Sufficient primary features remain for igneous trends to be studied in considerable detail.

The complex consists dominantly of layered cumulates, of which a total thickness of ~10 km occurs in the intrusion as a whole. Cryptic layering reveals reversals to higher-temperature mineral compositions with increasing height in the cumulate stratigraphic succession, which is believed to reflect repeated magma influx¹². Despite these reversals, the highest temperature mineral assemblage in the complex occurs at the base (olivine Fo₈₆ + picotite) and the lowest temperature assemblage occurs at the top in quartz-bearing ferrosyenites (albite, hedenbergite, ferro-edenite, ilmenite, apatite, quartz, K-feldspar, zircon, biotite and allanite). These syenitic differentiates, which grade downwards into layered ferrogabbroic cumulates, occupy <3% of the exposed area of the complex. Taking the entire stratigraphic sequence, olivine spans the range Fo₈₆ to Fo₀ with a hiatus between about Fo₇₁ and Fo₆₁, and plagioclase covers the range from An₈₀ to An_{1.5}. Ca-rich pyroxenes vary from Wo₄₅En₄₄Fs₁₁ to Wo₄₇En₀Fs₅₃ and Ca-poor pyroxenes span the range Wo₂En₆₇Fs₃₁ to Wo₂En₁₇Fs₈₁. Inverted pigeonite first occurs at Fs₆₇ (calculated on a Ca-free basis) and occurs sporadically together with primary orthopyroxene at compositions more iron-rich than this. Apparently magmatic grunerite occurs instead of Ca-poor pyroxenes in the later differentiates and reaches an Mg:Fe ratio of 12:88. Magnetite and ilmenite crystallize over wide temperature intervals. Calcic amphibole is an intercumulus phase throughout most of the complex, becoming a liquidus phase in the upper parts, as well as in sporadic layers in the lower part of the cumulate succession. Its composition ranges from titanian pargasite to ferro-edenite. Biotite is a liquidus phase in the most fractionated rocks.

The coexistence of two pyroxenes, the quartz-bearing nature of the final differentiates and the 'olivine gap' broadly indicate a

tholeiitic parental magma. The detailed igneous petrology of the complex will be published elsewhere.

Broad beam microprobe analyses, ensuring reliable bulk compositions, were made of 30 Ca-rich and 20 Ca-poor pyroxenes representing the entire cumulate succession. Analyses of the most Mg-rich and Fe-rich pyroxenes are given in Table 1. While there is considerable systematic variation in the minor elements Ti, Al, Mn and Na the major trend of crystallization, represented in Fig. 1, is one of increasing Fe:Mg ratio and relatively constant Ca for both Ca-rich and Ca-poor pyroxenes.

The Fongen-Hyllingen Ca-rich pyroxene trend is close to the calcic augite-calcic ferroaugite-hedenbergite trend of the Shiant Isles sill clinopyroxenes¹, but the latter, which crystallized from a mildly alkaline basalt parent, do not coexist with Ca-poor pyroxenes. Clinopyroxenes of strongly alkaline basic magmas, such as those of the Shonkin Sag laccolith³, are enriched in Na and Fe³⁺ in the later stages and consequently their evolution, which is not ideally expressed in a En-Di-Hd-Fs diagram, bears no comparison with our pyroxenes. The Fongen-Hyllingen two pyroxene trend is very different in character from the Skærgaard solidus trend^{4,5} (augite-ferroaugite-ferrohedenbergite with coexisting bronzite-inverted pigeonite) in the calcic nature of the Ca-rich pyroxenes, the extent of Fe-enrichment of the Ca-poor pyroxenes, and the very late appearance of inverted pigeonite. However, the Fongen-Hyllingen pyroxenes are very close to the Skærgaard subsolidus pyroxene trends¹³. The latter trends are based on microprobe point analyses of individual pyroxene exsolution lamellae/hosts, while our analyses were carried out using a defocused beam (diameter ~20 µm; exsolution lamellae on the micrometre scale) to provide solidus

Table 1 Broad-beam microprobe analyses of the most Mg and Fe-rich pyroxene compositions observed in the Fongen-Hyllingen complex

Analysis no.	Ca-rich pyroxene		Ca-poor pyroxene	
	A	B	C	D
SiO ₂	49.58	48.73	53.37	47.29
TiO ₂	1.03	†	0.26	0.11
Al ₂ O ₃	5.07	0.26	2.97	0.42
FeO*	7.10	29.75	15.33	45.40
MnO	0.15	0.78	0.35	1.62
MgO	15.20	†	26.63	5.34
CaO	21.55	20.63	1.13	0.90
Na ₂ O	0.50	0.35	†	†
Total	100.18	100.50	100.04	101.08

* All Fe expressed as FeO.

† Below detection limit.

compositions. This similarity is, therefore, more apparent than real and the Fongen-Hyllingen pyroxene compositions are consistent with crystallization under elevated P_{H_2O} , as indicated by the widespread occurrence of primary hydrous mafic phases.

The Fongen-Hyllingen pyroxene compositions form extensions of the calc-alkaline pyroxene trends of the Guadalupe complex^{7,8}. Guadalupe Ca-rich pyroxenes are restricted to the augite field and span the range Wo₄₃En₄₅Fs₁₂ to Wo₄₃En₃₀Fs₂₇, while the Ca-poor pyroxenes, which show no inversion from pigeonite, extend from Wo₃En₆₉Fs₂₈ to Wo₄En₃₅Fs₆₁. As a result of crystallization under fairly high P_{H_2O} in calc-alkaline complexes, the pyroxene liquidus surface is depressed, producing Ca-rich clinopyroxenes and extending the field of primary orthopyroxene crystallization to iron-rich compositions. Elevated P_{H_2O} also results in calcic amphibole being a liquidus phase at an early stage of fractionation, crystallizing together with two pyroxenes until the anhydrous pyroxenes become unstable and first Ca-rich pyroxene and then Ca-poor pyroxene cease to crystallize¹⁴. Calcic amphibole, together with biotite, then crystallize as the only mafic minerals until the end of fractionation.

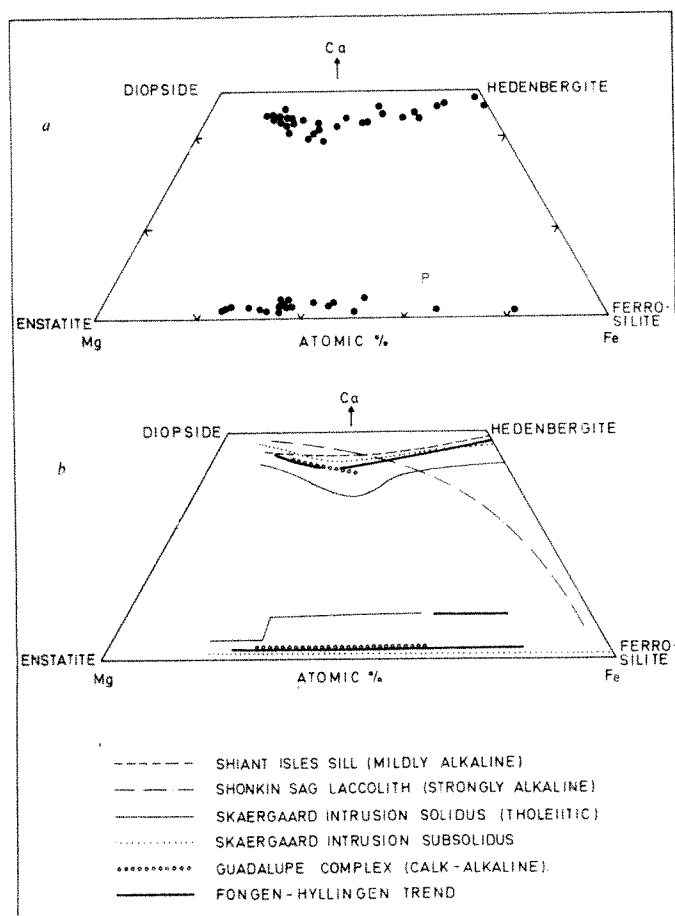


Fig. 1 *a*, Pyroxene compositions from the Fongen-Hyllingen complex. *P*, first occurrence of inverted pigeonite. No inverted pigeonites have been analysed. *P* is located in the pyroxene quadrilateral on tie line between coexisting calcic ferroaugite and ferrohypersthene. *b*, Comparison of the Fongen-Hyllingen pyroxene crystallization trend with those of other igneous complexes.

The extreme pyroxene trend presented here, with highly calcic Ca-rich pyroxene reaching Mg-free hedenbergite and a very extensive Ca-poor pyroxene trend with two Ca-poor pyroxenes coexisting at Fe-rich compositions, represents crystallization under moderately elevated and increasing P_{H_2O} . Water vapour pressure was higher than that typical of tholeiitic complexes but not high enough to produce the more restricted calc-alkaline trend. Other mineralogical features, such as the wide plagioclase and olivine fractionation ranges (reaching albite and pure fayalite respectively) and the occurrence of a small volume of syenitic differentiates as opposed to large volumes of granitic material, also distinguish the Fongen-Hyllingen fractionation from calc-alkaline complexes. We propose that this new pyroxene fractionation trend be named the Fongen-Hyllingen trend.

The Danish Natural Science Research Council and the Carlsberg Foundation provided financial support to study the Fongen-Hyllingen complex.

Received 17 November 1980; accepted 21 January 1981.

- Gibb, F. G. F. *J. Petrol.* **14**, 203 (1973).
- Aoki, K. *Am. Miner.* **49**, 1199 (1964).
- Nash, W. P. & Wilkinson, J. F. G. *Contr. Miner. Petrol.* **25**, 241 (1970).
- Brown, G. M. *Miner. Mag.* **31**, 511 (1957).
- Brown, G. M. & Vincent, E. A. *J. Petrol.* **4**, 175 (1963).
- Atkins, F. B. *J. Petrol.* **10**, 222 (1969).
- Best, M. G. *J. Petrol.* **4**, 223 (1963).
- Best, M. G. & Mercy, E. L. P. *Am. Miner.* **52**, 436 (1967).
- Nockolds, S. R. *J. geol. Soc. Lond.* **96**, 451 (1941).
- Wilson, J. R. & Olesen, N. Ø. *Norsk geol. Tidsskr.* **55**, 423 (1975).
- Olesen, N. Ø., Hansen, E. S., Kristensen, L. H. & Thyrsted, T. *Leid. geol. Meded.* **49**, 259 (1973).
- Esbensen, K. H., Thy, P. & Wilson, J. R. *Norsk geol. Tidsskr.* **58**, 103 (1978).
- Nwe, Y. Y. *Contr. Miner. Petrol.* **55**, 105 (1976).
- Cawthorn, R. G. *Am. Miner.* **61**, 907 (1976).

Submarine chutes on the slopes of fjord deltas

David B. Prior & Wm J. Wiseman Jr

Coastal Studies Institute, Louisiana State University, Baton Rouge, Louisiana 70803, USA

W. R. Bryant

Department of Oceanography, Texas A and M University, College Station, Texas 77843, USA

The first side-scan sonograph images of submarine landslide chutes from arctic fjords presented here indicate that submarine slopes of fjord head deltas are subject to instability. The chutes transport sediment from the upper delta slopes into deeper water parts of the fjords. Previously, sediment distribution and budgets in the bottom of deepwater fjords were ascribed to settling of suspended sediment supplied by streams and rivers during the short summer melt period. Previous descriptions of submarine mass movements from some subarctic fjords have been limited to isolated major catastrophic events. The sonographs from an arctic fjord suggest that delta-front instability occurs as multiple small-scale events that recur at frequent intervals and are intrinsic to the fjord sedimentation system.

The underwater topography of Adventfjord, a Spitsbergen fjord, and in particular actively prograding fluvioglacial deltas and their prodelta slopes, was examined using a side-scan sonar system. Chutes or delta-front gullies were distinctive features in some locations. Although the precise origins and functions of the chutes are obscure, morphological elements suggest that they may be due to underwater landsliding or mass movement processes; however, they are undoubtedly a significant component of some fjord delta geometries and associated progradational and depositional mechanisms. The present sonographs are from a small fan delta formed by active deposition from Longyear River into the side of Adventfjord and a major delta at the head of the fjord, the result of deposition from a braided outwash stream network, Advent River.

The submarine slopes in the fjord are steep. Water depths reach 30 m very close to the shoreline. Each delta has a well-developed planar bar formed by coalescence of adjacent distributary mouth bars, which produces a relatively flat feature composed of sand, silt, and fine gravel and covered by 1–3 m of water at low tide (tidal range 1–2 m). The delta-front slopes, seawards of the bars, are steeply inclined at angles of 12–16°. At the base of the delta slopes the fjord bottom is essentially flat at water depths of 60 m. The deltas demonstrate classical slope geometry and sediment lithology with fine sands on the bars and upper slopes and grain size becoming progressively finer seaward.

Chutes were observed in the upper slopes of the deltas between 5 and 30-m water depth, that is, where bottom slope angles are greatest. Typically the chutes begin on the steepest parts of the delta on the seaward margin of constructional aprons or bars, and are not incised into the nearshore surfaces of the bars. Thus, the chutes are not simply seaward extensions of distributary channels over the top of the submerged bars, but rather originate in arcuate re-entrants, with steep head slopes, cut into the bar fronts. The chutes usually exhibit an upslope bowl-shaped source region with an upslope scarp, an elongated narrow gully section, and a downslope extensive hummocky zone.

Figure 1 shows a sonograph and interpretive map of a system of chutes radiating out from the bar front of the Longyear fan delta. The bar is extremely irregular, with some large scarped and stepped areas. The scarps occasionally seem to bound

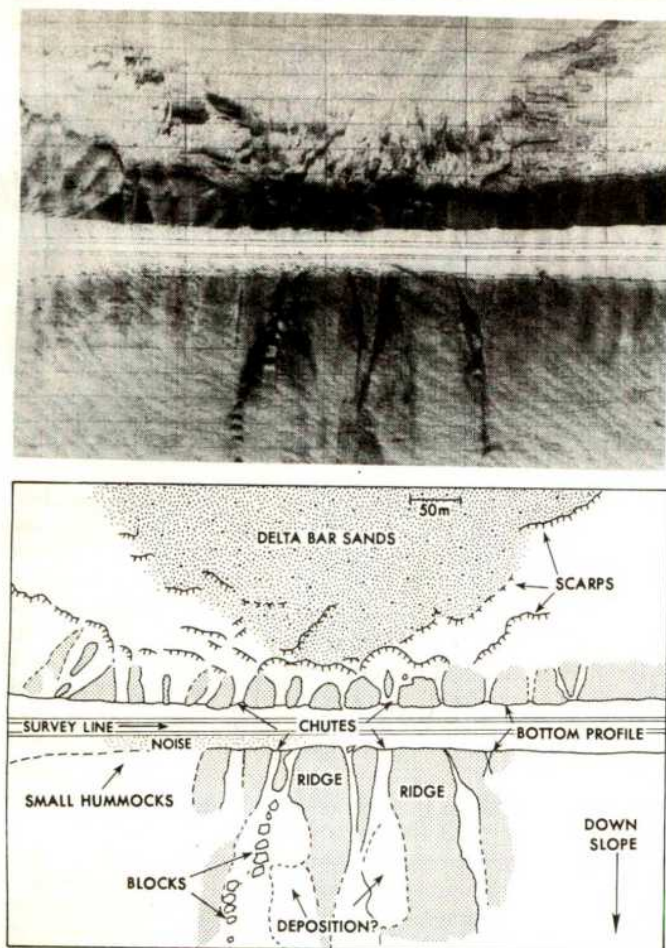


Fig. 1 Side-scan sonograph of Longyear delta front and interpretive map.

masses that have subsided and moved seawards. A fairly continuous scarp, extending around the front of the delta, shows major re-entrants immediately above distinct chutes. The chutes begin as bowl-shaped subsidence areas but become narrower farther downslope. The subsidence areas are typically 20–25 m across, whereas the elongate chutes average 10 m in width, are incised into the delta slope to depths of 2–5 m, and have no identifiable associated levees. The chutes show two distinct tonal patterns on this sonograph. In the central part of the delta, five chutes show very dark grey tones, whereas residual interchute ridges have relatively lighter grey tones. In contrast, chutes towards the edges of the fan are lighter toned than the intervening ridges. The precise reasons for such striking tonality differences are unknown but may be due to relative age, the darker toned chutes being relatively younger with incision exposing coarser-textured delta-front deposits, the lighter toned chutes being partially infilled by fine-grained deposits settling from suspension. Figure 1 also reveals that the lower (downslope) ends of the five freshest features are marked by irregularly shaped depositional areas of blocks and spreads of dark-toned sediments.

Figure 2 shows a series of chutes on the braided outwash delta at the head of Adventfjord. Chutes again emanate from bowl-shaped source areas backed by scarps on the edge of the distal bar. The chutes are represented by lighter toned areas separated by dark-toned interchute ridges and intact platforms. They vary in width from 15 to 30 m. In some cases adjacent chutes coalesce. The incisions are relatively shallow (<3 m) where the survey line crosses the chutes. Downslope from the chutes the delta-front slope is very hummocky, with contrasting light and dark tones, indicating considerable local relief over a large area. At one place, where several chutes converge, the downslope

area consists of several subparallel blocks and scarps. Other survey data show that the chutes are present across the entire bar front and that the middle delta slope is hummocky. The hummocks, blocks, and scarps decrease in size towards the foot of the delta slope, finally merging with the flat fjord floor.

The origin of the features in Figs 1 and 2 is not precisely known. However, an attractive hypothesis is that the bowl sources, chutes, and hummocky slopes are due to submarine sediment mass movement on the actively prograding delta slopes. Mass movement (slope instability) is an important process in many other active delta environments. For example, detailed sonar mapping in the Mississippi Delta region has revealed similar mudslide gullies^{1–6}. Other examples include the Copper River delta (Alaska)⁷, the Fraser Delta (British Columbia)⁸, and the Esmeraldas and Magdalena deltas (South America)^{9,10}. There is increasing evidence of sea floor sediment mass movement specifically in subarctic fjords and glacier-fed lacustrine delta areas. For example, one classic fan delta-front failure in Howe Sound fjord (British Columbia) caused extensive damage to adjacent buildings¹¹, as did slide-generated waves in Kitimat Inlet (British Columbia)¹². A survey of the Howe Sound delta with side-scan sonar showed the presence of delta-front chutes¹³. Similarly, the geological interpretation of relict and ancient fjord deltaic sediments is relying more and more on mass movement concepts and processes¹⁴.

The combination of circumstances in fjord deltas are conducive to slope failure. Bottom slopes are steeply inclined, and rapid progradation, especially during spring melt, causes temporary oversteepening and increases gravitational stress. Rapid deposition of sands, silts, and clays can cause local excess pore pressures, especially where fine sediment layers impede pore fluid drainage. Furthermore, bar sands are built out over previously deposited fjord bottom clays and glacial flour; the latter may not be able to accept the prograding bar loads, and

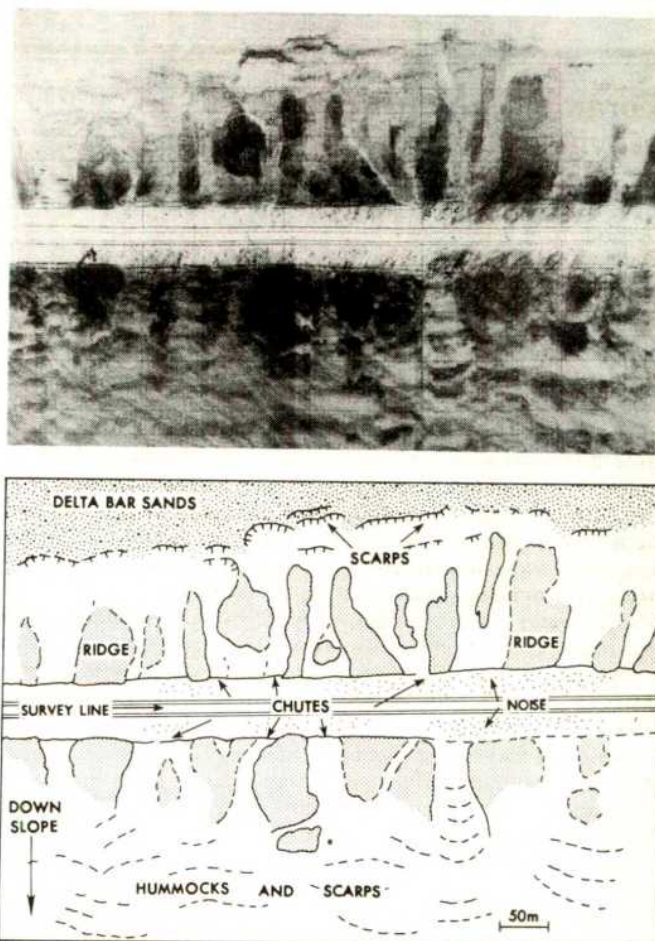


Fig. 2 Side-scan sonograph of Adventfjord delta front and interpretive map.

failures may result. Finally, upfjord winds can cause surface waves, which, as they shoal over the edge of the bar, may impart significant cyclic loading effects on the bar-front sediments.

The chutes and subsidence areas on delta slopes in Spitsbergen can be interpreted as the result of avalanching or debris flow movement of the bar-front sediments. A cyclic wave-induced loading of the sediments can generate pore fluid pressures sufficient to liquefy the bar-front deposits, which are already at or near the submerged angle of repose due to normal rapid sedimentation. Similarly, locally rapid depositional oversteepening could have the same effect. The transport of sediments down the chutes transfers loading stresses to the middle slopes of the delta, which may then, in turn, suffer failure, producing irregular hummocky terrain.

As similar side-scan sonar surveys are made for other fjord deltas, it is anticipated that delta-front chutes and other results of mass movement processes may prove to be extremely common.

This study was sponsored by the Arctic Program and Coastal Science Program of the Office of Naval Research, Arlington, Virginia.

Received 14 October 1980; accepted 23 January 1981.

1. Prior, D. B. & Coleman, J. M. *Geoscience and Man* Vol. 19 41–53 (Louisiana State University, Baton Rouge, 1978).
2. Prior, D. B. & Coleman, J. M. *Mar. Geotech.* **3**, 37–60 (1978).
3. Prior, D. B., Coleman, J. M. & Garrison, L. E. *Geology* **7**, 423–425 (1979).
4. Prior, D. B., Coleman, J. M., Suhayda, J. N. & Garrison, L. E. *Proc. Conf. on Port and Ocean Engineering under Arctic Conditions*, Trondheim, **2**, 921–933 (1979).
5. Prior, D. B. & Coleman, J. M. *Mar. Geol.* **36**, 227–239 (1980).
6. Coleman, J. M., Prior, D. B. & Garrison, L. E. *Bureau of Land Management Rep.* 80–81 (New Orleans, 1980).
7. Hampton, M. A. *et al. Offshore Tech. Conf.* Houston, Texas, Pap. 3314, 2307–2318 (1978).
8. Mathews, W. H. & Shepard, F. P. *Bull. Am. Ass. petrol. Geol.* **46**, 1416–1437 (1962).
9. Shepard, F. P. *Trans. Am. geophys. Un.* **13**, 226–230 (1932).
10. Heezen, B. C. *Bol. Soc. Geogr. Colombia* **51–52**, 135–143 (1956).
11. Terzaghi, K. *Proc. 8th Texas Oil Mech. Engr. Conf.* 1–41 (1956).
12. Murty, T. S. *J. geophys. Res.* **84** (C12), 7777–7779 (1979).
13. Prior, D. B., Wiseman, Wm. J. & Gilbert, R. (in preparation).
14. Shaw, J. *Geogr. Annlr* **59** (A), 221–240 (1977).

Foraminifera distribution of provinces in the Gulf of Mexico

Stephen J. Culver & Martin A. Buzas

Department of Paleobiology, Smithsonian Institution, Washington DC 20560, USA

Recent benthic foraminifera distribution patterns form the basis of much palaeoenvironmental (particularly palaeobathymetric) interpretation used by geologists. We have recently summarized the distribution of recent benthic foraminifera on the Atlantic continental margin of North America^{1,2} and report here the results of an analysis of all published data from the Gulf of Mexico. To obtain a synthesis of foraminiferal distribution in the Gulf of Mexico, a computerized catalogue of all published occurrences (presence or absence data) was compiled. Cluster analysis of these data isolated four large, marginally overlapping areas or provinces exhibiting a spatial correlation with Gulf of Mexico water masses. A fifth, smaller, discrete area (biofacies) was found at the mouth of the Mississippi delta.

There have been 77 papers in the past 60 years describing the distribution of recent benthic foraminifera in the Gulf of Mexico. We compiled this information using a large computer system. Our files record the distribution of 1,219 species names from 426 localities³. Through synonymization using published illustrations and specimens deposited in the US National Museum and the British Museum, the number of species was reduced to 848. Because cluster analysis of commonly occurring species on the Atlantic continental margin produced comparable results to analysis involving all species⁴, we used the 295 most commonly recorded species (those occurring at 4% or more of the 426 localities) for cluster analysis of the Gulf of Mexico data. Due to computer limitations the number of locali-

ties involved in analysis was reduced randomly to 350. The data matrix was analysed in the *Q*-mode by clustering Jaccard coefficients by the weighted pair-group method^{5–7}.

The cluster analysis produced four large, marginally overlapping areas or provinces (Fig. 1): the Coastal, Inner Shelf, Outer Shelf, and Slope and Abyssal Plain Provinces. (The diagnostic species associations of the provinces will be published elsewhere.) A fifth, smaller, discrete area or biofacies is located at the mouth of the Mississippi River.

Our analysis, based on all published presence or absence data for 295 commonly recorded species of benthic foraminifera, indicate that for the sampled littoral, inner shelf and outer shelf areas, the same foraminifera assemblages occur all around the Gulf and comprise the Gulf of Mexico Coastal, Inner Shelf and Outer Shelf Provinces respectively (Fig. 1). In deeper water, the Gulf of Mexico Slope and Abyssal Plain Province covers the whole sampled sea bed beneath depths of ~1,000 m (Fig. 1). Our circum-Gulf provinces agree with the pattern of benthic foraminifera distributions reported by Poag⁸, but he divided the deep basin into distinct eastern and western assemblages as did Dignes²⁰.

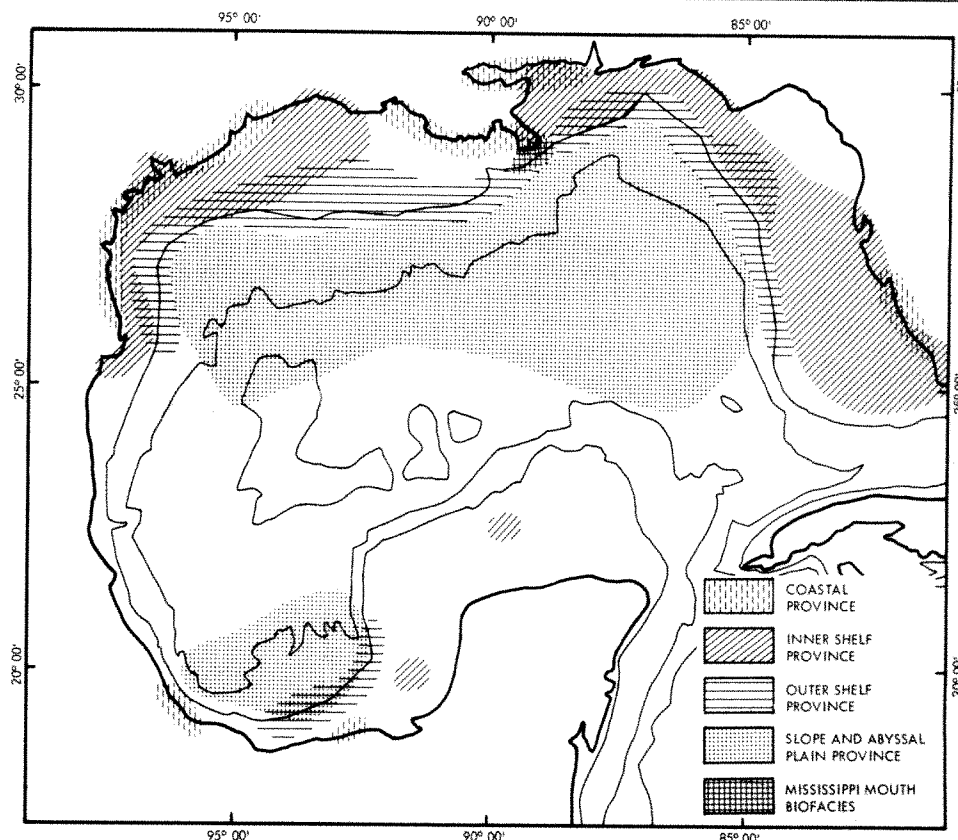
The provinces defined by our analysis seem to be unrelated to sediment substrate. For example, the Yucatan and west Florida carbonate shelves and the clastic northern Gulf of Mexico shelf east and west of the Mississippi delta are all included in the Gulf Inner Shelf Province. In agreement with macrofaunal patterns⁹, therefore, foraminifera shelf faunas on both sides of the Mississippi delta are generally the same. However, at the mouth of the delta, 10 localities group together as a discrete biofacies (Fig. 1). This group is considered to be too small and geographically restricted to be referred to as a province, but its low diversity assemblage shows close affinities to the Outer Shelf Province assemblage. The position of this biofacies is probably related to oceanographic or sedimentological conditions related to the Mississippi River discharge.

Various provincial schemes have been proposed for the Gulf of Mexico–Caribbean region using shallow water macro-invertebrates and fishes^{9–15}. The northern Gulf coast from Cape Romano, Florida to approximately the Mexico–USA border is now generally considered to be part of the east coast of North America Carolinian Province^{9,14,15}. The southern Gulf coast and the southern tip of Florida from Cape Romano to Cape Canaveral is part of the Caribbean Province, which includes the West Indian islands and the northern coast of South America^{9,14,15}. Other schemes involve greater subdivision^{10,13} or the inclusion of all the Gulf of Mexico into a Caribbean Province¹².

We compared the shallow-water benthic foraminifera assemblages from the east coast of North America south of Cape Hatteras⁴ with shallow water assemblages in the Gulf of Mexico (data unpublished). An inner continental shelf Carolinian Province, similar in extent to that defined by macro-invertebrate assemblages (see above), does not exist for benthic foraminifera. That is, inner shelf foraminifera faunas of the northern Gulf of Mexico differ considerably from those of the shelf south of Cape Hatteras. Sixty commonly recorded species are present on the east coast Southern Shelf Province, and 172 are present in the Gulf of Mexico Inner Shelf Province; only 42 species (22% of a total of 189 species) are commonly recorded species in both provinces. The question remains whether the Gulf of Mexico, based on foraminifera data, is part of a larger Caribbean Province.

A congruous relationship between foraminifera provinces and bottom water masses in the Gulf of Mexico is demonstrable. As for the Atlantic Coastal Province¹, we suggest the Gulf of Mexico Coastal Province is related to the marginal waters of bays, estuaries and lagoons. The Inner and Outer Shelf Provinces show varying amounts of overlap. The distinction between these provinces can be related to the difference in magnitude of seasonal effects between the inner and outer shelf. The critical limits of the set of environmental variables affecting the foraminifera assemblages will be located at different

Fig. 1 Recent benthic foraminifera provinces in the Gulf of Mexico; 183 m, 1,830 m, 3,660 m submarine contours are shown.



distances from the shore from year to year (or season to season); hence the considerable degree of overlap. The Inner Shelf Province on the southern margin of the Gulf may be less well defined due to sampling deficiencies or to decreased environmental variability, in particular less severe salinity changes due to a relative lack of river discharge compared with the northern margin of the Gulf.

The depth limit of seasonal effects on water mass characteristics is ~100–200 m (refs 21, 22). This depth correlates with the basinward limit of the Inner Shelf Province. Although the open surface waters of the Gulf can be divided into eastern and western parts¹⁶, the deeper water, below ~1,000–1,500 m, has generally uniform temperature, salinity and dissolved oxygen relations^{16–19}. Thus, the disposition of the pan-Gulf of Mexico Slope and Abyssal Plain Province (Fig. 1) agrees with that of the bottom water mass (that is, no division into eastern and western provinces) and an upper limit at ~1,000 m. In agreement with the provincial pattern presented here, distribution maps for Gulf of Mexico commonly recorded species³ show that all species, except one, that occur at abyssal depths range up on to the slope or even the outer shelf. Thus, there is no solely abyssal association composed of species restricted to abyssal depths.

We thank C. G. Adams, L. Isham, A. Lanham, C. McCloy, K. McCormick, D. Piacesi, D. Schnitker, D. J. Stanley, R. Cifelli and J. Whittaker for assistance. The research was supported by a Smithsonian Scholarly Studies Award.

Received 11 December 1980; accepted 3 February 1981.

1. Buzas, M.A. & Culver, S.J. *Science* **209**, 687–689 (1980).
2. Culver, S. J. & Buzas, M. A. *Smithson. Contr. Mar. Sci.* **6**, 1–512 (1980).
3. Culver, S. J. & Buzas, M. A. *Smithson. Contr. Mar. Sci.* (in the press).
4. Culver, S. J. & Buzas, M. A. *J. Foram. Res.* (in the press).
5. Sokal, R. R. & Sneath, P.H.A. *Principles of Numerical Taxonomy* (Freeman, San Francisco, 1963).
6. Kaesler, R. L. *Kans. Paleontol. Contr.* **10**, 1–50 (1966).
7. Mello, J. F. & Buzas, M. A. *J. Paleontol.* **42**, 747–758 (1968).
8. Poag, C. W. *Abstr. Benthonics '75 '35* (Dalhousie University, Nova Scotia, 1975).
9. Hedgpeth, J. W. *Inst. Mar. Sci., Univ. Texas* **3**, 111–224 (1953); *Fish. Bull.* **89**, *Fish Wildlife Serv.* **55**, 203–214 (1954).
10. Briggs, J. C. *Marine Zoogeography* (McGraw-Hill, New York, 1974).
11. Fischer, P. *Manuel de Conchyliologie* (Savy, Paris, 1887).
12. Hall, C. A. *Ecology* **45**, 226–234 (1964).
13. Pulley, T. E. A. *Rep., Am. Malacol. Un.* 1952, 2–3 (1953).
14. Rehder, H. A. *Fish. Bull.* **89**, *Fish Wildlife Serv.* **55**, 469–474 (1954).
15. Warmke, G. L. & Abbott, R. T. *Caribbean Seashells* (Livingston, Narberth, 1961).
16. Ichiye, T. *Geofis. Int.* **2**, 47–76 (1962).

17. Harding, J. L. & Nowlin, W. D. Jr *Encyclopedia of Oceanography* (ed. Fairbridge, R. W.) 324–331 (Rheinhold, New York, 1966).
18. McLellan H. J. & Nowlin, W. D. Jr *J. Mar. Res.* **21**, 233–245 (1963).
19. Nowlin, W. D. Jr *Contributions on the Physical Oceanography of the Gulf of Mexico* (eds Capurro, L.R.A. & Reid, J.L.) 3–51 (Gulf, Houston, 1972).
20. Dignes, T. W. thesis, Univ. Maine (1978).
21. Nowlin, W. D. Jr & Parker, C. A. *J. phys. Oceanogr.* **4**, 467–486 (1974).

Mimicry of antenna and photo-protective carotenoid functions by a synthetic carotenoporphyrin

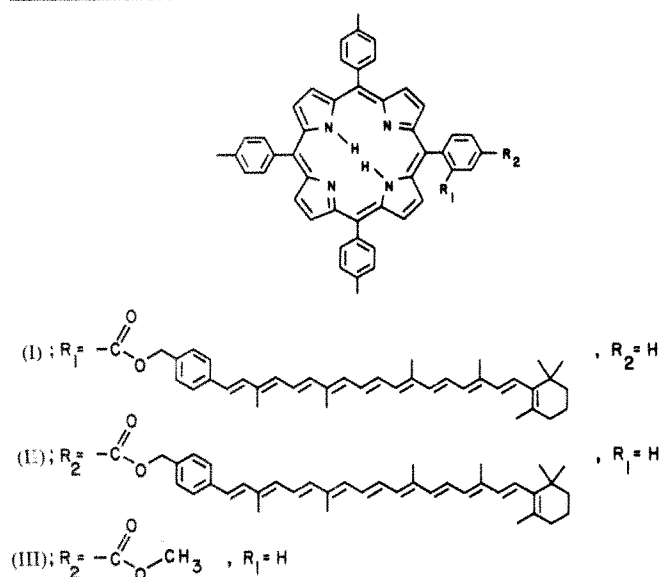
René V. Bensasson*, Edward J. Land†, Ana L. Moore‡, Robert L. Crouch‡, Gary Dirks‡, Thomas A. Moore‡ & Devens Gust‡

* Laboratoire de Biophysique, INSERM U201, Muséum National d'Histoire Naturelle, 61 Rue Buffon, 75005 Paris, France

† Paterson Laboratories, Christie Hospital and Holt Radium Institute, Manchester M20 9BX, UK

‡ Department of Chemistry, Arizona State University, Tempe, Arizona 85281, USA

Carotenoid pigments, ubiquitous in photosynthetic membranes, are essential for the survival of green plants¹. Two facets of carotenoid function are recognized in photosynthetic membranes. First, carotenoids prevent the chlorophyll-photo-sensitized formation of highly destructive singlet oxygen by intercepting the chlorophyll triplet states^{2–10} and may also scavenge additional singlet oxygen present^{11,12}. Second, carotenoids perform an antenna function by transferring the energy of absorbed light at the singlet excited state level to the chlorophyll system for the execution of photosynthetic work^{13–16}. Nevertheless, the mechanisms by which carotenoids perform these functions are poorly understood. We now report that a unique synthetic carotenoporphyrin I consisting of a carotenoid part covalently linked to a synthetic tetraarylporphyrin successfully mimics both the photophysical functions of carotenoids in photosynthesis. The explanation for this seems to be the close interaction of the carotenoid and porphyrin π -electron systems.

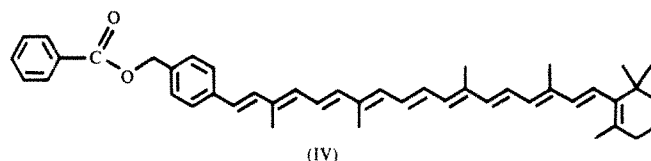


We have previously shown that singlet-singlet energy transfer from the carotenoid moiety of carotenoporphyrin II to the porphyrin is vanishingly small¹⁷. Thus, simply placing a carotenoid in the proximity of a porphyrin is not enough to ensure antenna function. In I, on the other hand, carotenoid to porphyrin singlet-singlet energy transfer occurs with 25% efficiency¹⁸. The enhanced energy transfer efficiency in I is due to the fact that unlike II, I adopts a folded conformation in which the carotenoid π -electron system lies directly over that of the porphyrin (Fig. 1). The close approach of the two chromophores suggests that an electron exchange interaction may contribute to the observed singlet-singlet energy transfer¹⁹. In view of the fact that triplet-triplet energy transfer is necessarily mediated through exchange terms²⁰, and because singlet oxygen arises from the interaction of oxygen with a sensitizer triplet², we have investigated the photochemical and dynamic properties of the triplet states of I and II.

Singlet oxygen can be conveniently detected by spectroscopically monitoring its reaction with diphenylisobenzofuran (DPBF)²¹. In Fig. 2, singlet oxygen production detected in this manner is plotted for a variety of sensitizers including I, II, and a simple tetraarylporphyrin, III. In addition, the effect of added

carotenoids on photoproduction of singlet oxygen by III is shown. The most striking feature is the essentially complete photoprotection afforded by I when contrasted with the extensive singlet oxygen production by II. If efficient photoprotection were simply a matter of relative concentration with no specific geometric requirements, one might expect similar degrees of photoprotection in I and II. This is most definitely not the case, and an understanding of the results required a more detailed investigation of the photophysics of these systems.

Consequently, we have carried out pulse radiolysis and laser flash experiments to monitor directly the evolution of the triplet states²². The properties of the carotenoid triplet states were established by pulse radiolysis. Population of the lowest triplet states of I, II and IV by pulse radiolysis in benzene through sensitization by biphenyl²² resulted in a transient spectrum with $\lambda_{\text{max}} = 535 \text{ nm}$ and $\epsilon_T - \epsilon_G \sim 1.4 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$, which is characteristic of a carotenoid triplet^{23,24} (ϵ_T and ϵ_G are the extinction coefficients of triplet and ground states, respectively). Similar experiments performed with III resulted in an $\epsilon_T - \epsilon_G \sim 6.8 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ at 440 nm, the transient difference spectrum being typical of a porphyrin. It is important to note that these transient spectra observed through pulse radiolysis for I, II and



IV are the same: the spectra and extinction coefficients for I and II are not perturbed by the covalent linkage between the two chromophores.

We now turn to the laser flash experiments. Figure 3 presents oscilloscope traces of the transient absorption changes with time for II after 353 nm laser excitation. The porphyrin triplet observed at 440 nm (Fig. 3b) appears first with $t_{\frac{1}{2}} \sim 1.5 \mu\text{s}$. Concomitant with this decay, and in a 1:1 ratio, the carotenoid triplet observed at 535 nm (Fig. 3a) rises with the same $t_{\frac{1}{2}}$. Subsequently, this absorption decays with $t_{\frac{1}{2}} \sim 7 \mu\text{s}$ (Fig. 3c). Taking into account the carotenoporphyrin concentration and the maximum possible expected intermolecular triplet energy transfer rate, this result suggests that the carotenoid is quenching the porphyrin triplet state by an intramolecular energy

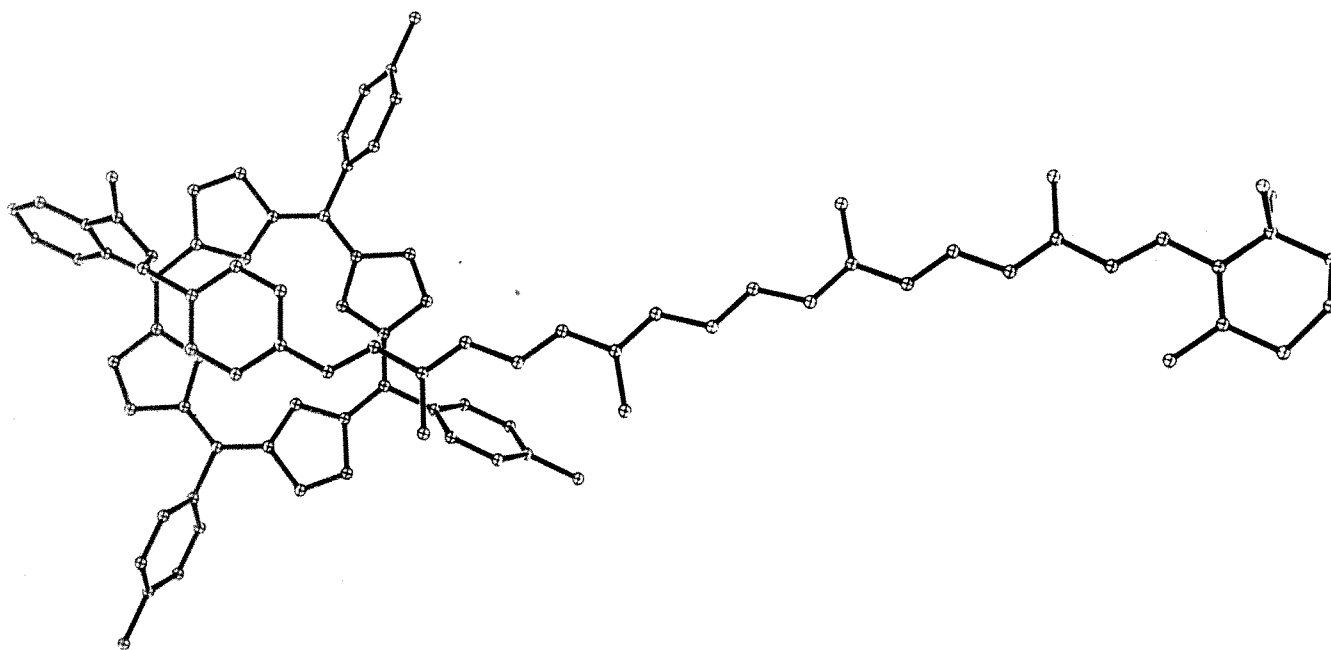


Fig. 1 Idealized time average solution conformation for carotenoporphyrin I. The aromatic ring of the carotenoid moiety is $\sim 4 \text{ \AA}$ above the mean porphyrin plane. Hydrogen atoms are not shown. This conformation is based on proton NMR spectral data¹⁸.

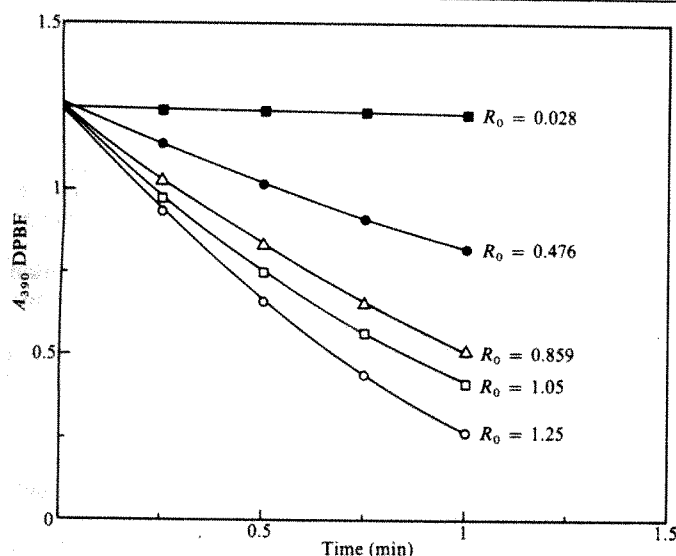


Fig. 2 Bleaching of DPBF at 390 nm by singlet molecular oxygen. Filtered light (630 nm long pass cutoff, Oriel G-772-6300 and a water filter) from a 150-W tungsten halogen projection lamp was used to illuminate the solutions. Reactions were carried out in spectrophotometer cells open to the air at solution volumes of ~2 ml. Sample temperature was $23 \pm 1^\circ\text{C}$. All solutions were prepared from freshly distilled toluene stored over molecular sieves. All compounds were purified by TLC immediately before use except DPBF which was used as received (Sigma). The DPBF concentration was $2.1 \times 10^{-6}\text{ M}$ in all experiments. ■, $1.9 \times 10^{-6}\text{ M}$ I; ●, $1.8 \times 10^{-6}\text{ M}$ II; △, $1.9 \times 10^{-6}\text{ M}$ III plus $1.9 \times 10^{-6}\text{ M}$ β -carotene; □, $1.9 \times 10^{-6}\text{ M}$ III plus $1.9 \times 10^{-6}\text{ M}$ benzoate ester of 7'-apo-7'-(4-hydroxymethylphenyl)- β -carotene (IV); ○, $2.0 \times 10^{-6}\text{ M}$ III. The wavelength monitored (390 nm) was chosen for its relative freedom from absorbance by chromophores other than DPBF. At each point the entire absorption spectrum (200–800 nm) was acquired (Hewlett-Packard 8450 A spectrometer) to ensure that no other photochemical processes were occurring. Plotted absorbance values have been corrected by subtracting residual absorbance at 390 nm of the solution without DPBF (typically A_{390} 0.06–0.09). The initial rates (R_0) are $(A_{t_0} - A_{t_{15}})/15\text{ s}$. Illumination of $2.1 \times 10^{-6}\text{ M}$ DPBF showed the same response as I within experimental error. There was no change in absorbance at 390 nm when $2.0 \times 10^{-6}\text{ M}$ I was illuminated.

transfer process. Figure 4 presents the results for a similar experiment for I. The behaviour of this carotenoporphyrin is strikingly different. The carotenoid triplet appears immediately (Fig. 3a); the porphyrin triplet state cannot be observed even at the time resolution limit (~30 ns) of our instrument (Fig. 3b). The quantum yield, determined by the comparative method²², for the formation of the porphyrin triplet state in III was 0.45 whereas the quantum yield of carotenoid triplet state, and consequently also the yield for the formation of the porphyrin triplet state, was 0.23 in II and 0.12 in I. (These quantum yields are consistent with the increase in quenching of the porphyrin fluorescence from III to II to I^{17,18}. The competing nonradiative path from the porphyrin singlet state is of interest because electron transfer between the chromophores may be occurring. This mechanism has been proposed for the quenching of chlorophyll *a* fluorescence by β -carotene²⁵.)

From these results it is clear that the remarkable photoprotective function observed for I (see Fig. 2) is a consequence of the extremely rapid quenching of the porphyrin triplet state by the carotenoid moiety. *In vitro* studies of bacterial systems have shown that quenching of the bacteriochlorophyll triplet state by carotenoid occurs with $t_{1/2} \sim 20\text{ ns}$ (ref. 6). The difference in porphyrin triplet quenching between I and II shows that achieving the required quenching rates on the nanosecond time scale requires a conformation such as that found for I (Fig. 1), which allows a high degree of orbital overlap and therefore very strong exchange interactions. Such a conformation is not available for II.

It is interesting to speculate on the implications of the above results for the elucidation of the photoprotective and antenna functions of carotenoids *in vivo*. A stacked conformation with close approach of the π -electron systems was necessary in I both for antenna function in the singlet manifold (where energy transfer must compete with rapid internal conversion in the carotenoid)^{17,18,27} and for efficient suppression of singlet oxygen production via the triplet manifold (where porphyrin triplet quenching by the carotenoid must compete with quenching by abundant oxygen). Such a conformation was necessary because both processes require relatively strong electronic interactions between the carotenoid and porphyrin π -electron systems. Perhaps a similar arrangement of pigments, with a separation of only a few ångströms, is necessary in photosynthetic reaction centres. Indeed, results for I and II suggest that the photoprotective and antenna functions of a particular carotenoid molecule *in vivo* may involve the same chlorophyll molecule. Such intimate and specific pigment-pigment interactions are doubtless mediated by the higher-order structure of the pigment-protein complexes. Moreover, photoprotection requires a chromophore with a triplet energy below that of singlet oxygen (1 eV). Such an organic molecule must have an extensive π -electron system and will therefore strongly absorb visible light.

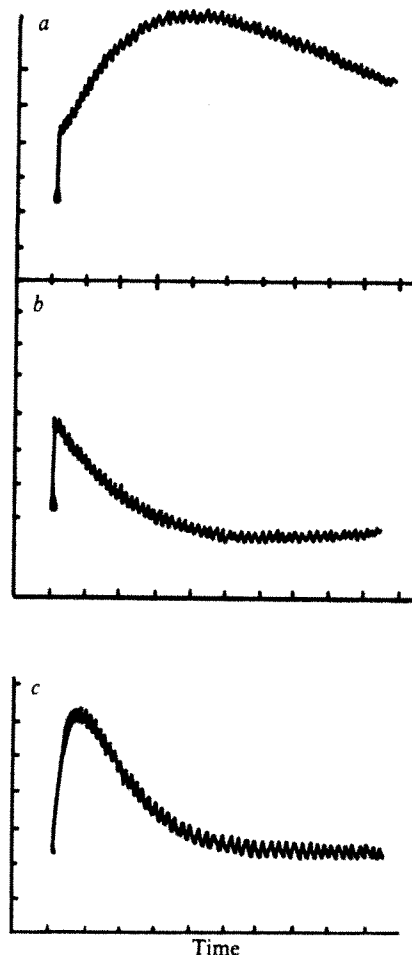


Fig. 3 Transient changes in absorbance of II after 353 nm laser flash. The flash was from a Q-switched, frequency tripled neodymium doped laser delivering 30 mJ with a half-peak-height pulse duration of 25 ns. The flash photolysis apparatus has been described previously²⁶. Samples were $\sim 4 \times 10^{-6}\text{ M}$ in benzene solution and were deaerated by bubbling argon. a, Carotenoid triplet absorption at 535 nm, 1 μs per div., 6.1% change in transmission (δT) per div.; b, Porphyrin triplet absorption at 440 nm, 1 μs per div., 6.8% δT per div.; c, carotenoid triplet absorption at 535 nm, 5 μs per div., 6.1% δT per div.

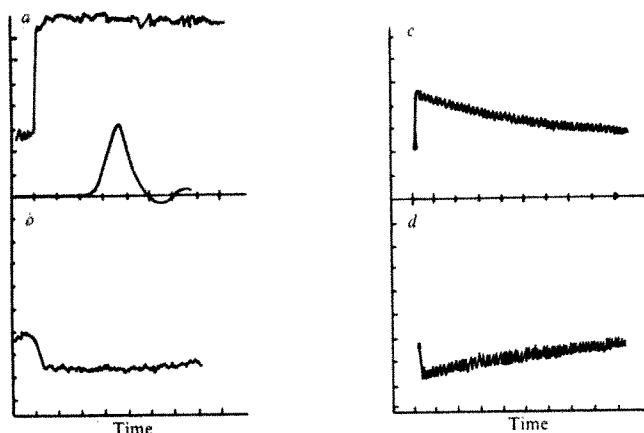


Fig. 4 Transient changes in absorbance of I after 353 nm laser flash. All conditions as in Fig. 3 legend. *a*, Carotenoid triplet absorption at 535 nm, 100 ns per div., 2.4% δT per div. (inset is laser pulse swept at 20 ns per div.); *b*, absorption at 440 nm, 100 ns per div., 2.8% δT per div. This change is due to depletion of the carotenoid ground state. *c*, Carotenoid triplet absorption at 535 nm, 1 μ s per div., 4.4% δT per div.; *d*, absorption at 440 nm, 1 μ s per div., 2.8% δT per div. This change is due to the repopulation of the carotenoid ground state.

The occurrence of a carotenoid-chlorophyll-protein ultra-structure which collects this light for photosynthesis while providing photoprotection is characteristic of the economy of biological systems.

Finally, biomimetic synthetic solar energy conversion systems using chlorophylls or porphyrins may well inherit problems with light-gathering efficiency and photoprotection similar to those encountered *in vivo*. As demonstrated above, carotenoid polyenes might provide a biomimetic solution to these problems.

This material was prepared with the partial support of the US Department of Energy, grant DE-FG02-79ER10545. This work was also supported by grants from the Cancer Research Campaign and the MRC. We thank Hoffmann-La Roche, Inc., for a gift of β -apo-8'-carotenal.

Received 29 October; accepted 15 December 1980.

- Goodwin, T. W. in *Chemistry and Biochemistry of Plant Pigments* Vol. 1 (ed. Goodwin, T. W.) 228 (Academic, London, 1976).
- Griffiths, M., Sistrom, W. R., Cohen-Bazire, G. & Stanier, R. Y. *Nature* **176**, 1211-1214 (1955).
- Cohen-Bazire, G. & Stanier, R. Y. *Nature* **181**, 250-252 (1958).
- Krinsky, N. I. in *Photophysiology* Vol. 3 (ed. Giese, A. C.) 123-195 (Academic, New York, 1968).
- Wolff, Ch. & Witt, H. T. Z. *Naturforsch.* **24b**, 1031-1037 (1969).
- Monger, T. G., Cogdell, R. J. & Parson, W. W. *Biochim. biophys. Acta* **449**, 136-153 (1976).
- Kung, M. C. & DeVault, D. *Photochem. Photobiol.* **24**, 87-91 (1976).
- Renger, G. & Wolff, Ch. *Biochim. biophys. Acta* **460**, 47-57 (1977).
- Boucher, F., Van Der Rest, M. & Gingras, G. *Biochim. biophys. Acta* **461**, 339-357 (1977).
- Mathis, P., Butler, W. L. & Satoh, K. *Photochem. Photobiol.* **30**, 603-614 (1979).
- Krinsky, N. I. in *Carotenoids* (eds Isler, O., Guttman, G. & Solms, U.) 669-716 (Birkhäuser, Basel, 1971).
- Foote, C. S., Chang, Y. C. & Denny, R. W. J. *Am. chem. Soc.* **92**, 5216-5217 (1970).
- Koka, P. & Song, P. S. *Biochim. biophys. Acta* **495**, 220-231 (1977).
- Bazzaz, M. B. & Govindjee, P. *Physiol.* **52**, 257-262 (1973).
- Goedheer, J. D. *Biochim. biophys. Acta* **172**, 252-265 (1969).
- Govindjee & Govindjee, R. in *Bioenergetics of Photosynthesis* (ed. Govindjee) 2-50 (Academic, New York, 1975).
- Dirks, G., Moore, A. L., Moore, T. A. & Gust, D. *Photochem. Photobiol.* **32**, 277-280 (1980).
- Moore, A. L., Dirks, G., Gust, D. & Moore, T. A. *Photochem. Photobiol.* **32**, 691-696 (1980).
- Razi Naqvi, K. *Photochem. Photobiol.* **31**, 523-524 (1980).
- Dexter, D. L. J. *Chem. Phys.* **21**, 836-850 (1953).
- Matheson, I. B. C. & Lee, J. *Chem. Phys. Lett.* **7**, 475-476 (1970).
- Bensasson, R. & Land, E. J. in *Photochemical and Photobiological Reviews* (ed. Smith, K.) Vol. 3, 163-191 (Plenum, New York, 1978).
- Bensasson, R., Dawe, E. A., Long, D. A. & Land, E. J. *JCS Faraday Trans. I* **73**, 1319-1325 (1977).
- Bensasson, R., Land, E. J. & Maudinas, B. *Photochem. Photobiol.* **22**, 189-193 (1976).
- Beddard, G. S., Davidson, R. S. & Tretheway, K. R. *Nature* **267**, 373-374 (1977).
- Bensasson, R., Chachaty, C., Land, E. J. & Salet, C. *Photochem. Photobiol.* **16**, 27-37 (1972).
- Dallinger, R. F., Woodruff, W. H. & Rogers, M. A. J. *Photochem. Photobiol.* **33**, 275-277 (1981).

Absence of ocular dominance patches in dark-reared cats

N. V. Swindale

Physiological Laboratory, Downing Street,
Cambridge CB2 3EG, UK

If a young monkey or kitten is monocularly deprived for a period of days or weeks, the ocular dominance stripes or patches formed in layer IV of the visual cortex by the geniculo-cortical afferents driven by that eye become smaller, while the patches formed by afferents from the other, experienced eye, spread out and increase in size^{1,2}. One explanation for this effect is that it results from a disturbance of a competitive process which, during the first weeks of life, guides a 'sorting out' of the initially intermixed right and left eye inputs into complementary, largely non-overlapping territories^{1,3}. One feature of this process may be a local interaction between right and left eye synapses in which like synapses reinforce each other's growth rates and cause rejection of the other eye's synapses^{1,4}. If this is the case, then the effect of monocular deprivation on the relative sizes of the two sets of columns can be explained by supposing that the strengths of the effects exerted by the deprived eye are reduced^{1,4}. This explanation has a testable consequence: if both eyes are deprived of vision then each eye should be made less effective in eliminating the other eye's inputs, and the overall rate at which the ocular dominance columns form should be decreased⁴. Although LeVay *et al.*⁵ found that columns were present in a 7-week old monkey reared in the dark from the age of 3 days, this result does not necessarily imply that the rate of column formation had been normal, because in normal monkeys the columns are well developed by 3 or more weeks of age⁵. I report here the results of transneuronal autoradiography in cats, which show that columns, as revealed anatomically, are undetectable in most parts of the visual cortex of cats reared in the dark for periods of up to 20 weeks, implying that visual experience is necessary for their proper formation.

Six kittens were used; four were reared in complete darkness and studied at the ages of 8, 10, 16 and 20 weeks, while two normal kittens of 8 and 10 weeks of age served as controls. Each animal had its right eye injected, under Halothane anaesthesia, with 2 mCi of ³H-proline dissolved in 100 μ l of saline. They survived a further 12-14 days, after which they were given a lethal dose of barbiturate and perfused with 4% paraformaldehyde in 0.1 M phosphate buffer. Frozen sections 30 μ m thick were cut in a horizontal plane through the visual cortex, mounted on gelatin-coated slides, defatted and dipped in Ilford K2 emulsion diluted 1:3 with 1% glycerol⁶. After exposure periods of 2 weeks to 2 months, the slides were developed in D19 (Kodak), fixed in hypo, and coverslipped without counterstaining.

Table 1 Degree of spillover in the lateral geniculate nuclei ipsilateral and contralateral to the injected eye

Age (weeks)	Right (ipsi)	Left (contra)
8 } Normal	16%	26%
10 } Normal	*	*
8 } Dark-reared	9%	21%
10 } Dark-reared	9%	17%
16 } Dark-reared	*	*
20 } Dark-reared	23%	12%

The table shows the degree of spillover in the lateral geniculate nuclei ipsilateral and contralateral to the injected eye (which was always on the right side) of the normal and experimental animals. The amount of label present in the lamina (A or A1) not innervated by the injected eye is expressed as a percentage of that in the lamina which was innervated. There is no suggestion that spillover is more severe in the dark-reared animals.

* Data unavailable.

As spillover of radioactive label into geniculate laminae not innervated by the injected eye has been shown to lead to an underestimate of the degree of segregation of the two eyes' terminals in the visual cortex³, I also examined the effect of dark rearing on this. The procedure was similar to that of LeVay *et al.*³: 1- μ m Araldite sections of the lateral geniculate nucleus were cut, coated with emulsion and processed in the same way as the frozen sections. Grain counts were made in laminae A and A1, which are innervated by the contralateral and ipsilateral eyes respectively, and the ratio of the two values expressed as a percentage (Table 1). No systematic differences between dark-reared and normal animals were found, although spillover was usually more severe on the contralateral side³.

The labelling pattern in both the 8- and 10-week old normal animals studied was similar to that found in normal adult animals^{3,7} (Fig. 1). In layer IV of area 17, ipsilateral to the injected eye, the label formed patches 200–600 μ m wide, spaced at roughly 800- μ m intervals, with a level of label in the gaps between which was only slightly higher than the background in adjacent layers of the cortex. In area 17 contralateral to the injected eye, the label was also distributed periodically, although the level of label in the gaps between the columns was higher than on the ipsilateral side. This probably results from the greater degree of spillover on this side.

All four dark-reared animals gave similar results. Over most of area 17, on sides both ipsilateral and contralateral to the injected eye, the distribution of label was almost uniform (Fig. 2), although in some parts of the cortex, slight fluctuations, whose periodicity was not obviously abnormal, were present. The level of label in the troughs of these fluctuations was always much higher than in the gaps between the labelled columns in the corresponding hemispheres of normal animals. They were always more pronounced in the ipsilateral hemisphere and were less noticeable or absent in regions of cortex receiving inputs

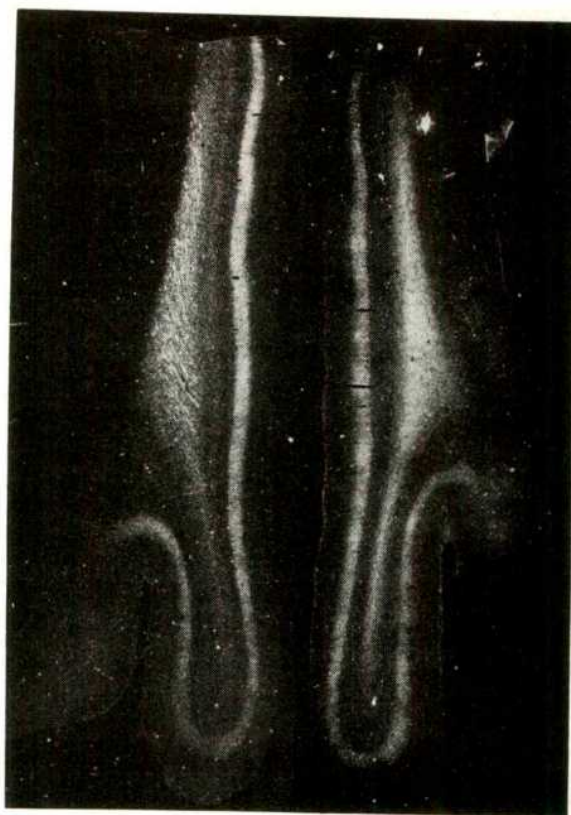


Fig. 2 Autoradiographs from the left and right hemispheres of an 8-week old dark-reared kitten, whose right eye was injected. The distribution of label is nearly uniform in both hemispheres.

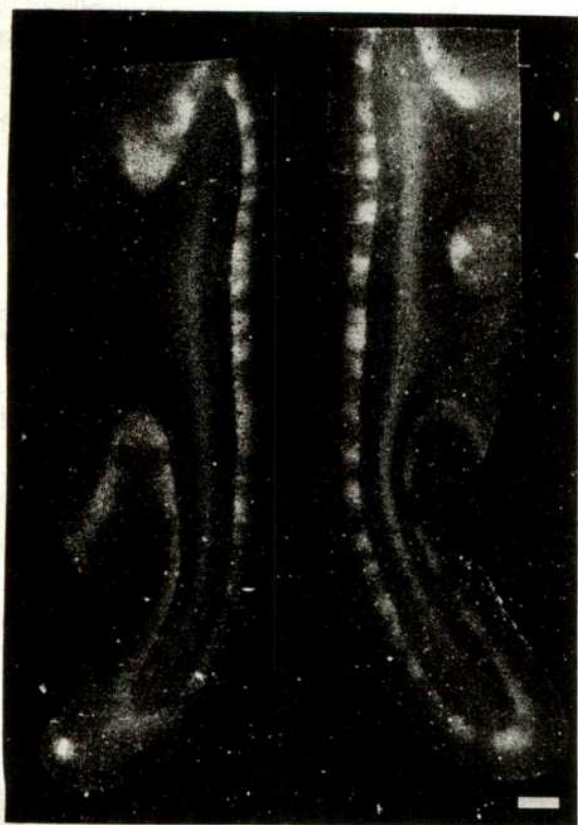


Fig. 1 Dark field autoradiographs from single horizontal sections through the left and right visual cortices of an 8-week old normal kitten whose right eye was injected with radioactive label. The label has a periodic, patchy distribution in layer IV of both hemispheres. In this and Fig. 2, the left hemisphere is on the left-hand side, and anterior is uppermost. Scale bar, 1 mm.

from peripheral retina. Contrary to expectation they were, if anything, less noticeable in the older dark-reared animals.

Although these results show that total visual deprivation interferes substantially with the process by which left and right eye terminals segregate to form ocular dominance columns, it remains unclear whether the deprivation is, as predicted, simply slowing the rate of segregation, in which case the columns should form eventually, or whether the process halts or even reverses after an intermediate stage of development. Some degree of development has occurred in the dark, because the fluctuations found in some parts of the cortex are absent from normal 2-week old animals³ (despite their several days' visual experience). The dissimilarity between these results from cats and those from the 7-week old dark-reared monkey studied by LeVay *et al.*⁵, which possessed an apparently normal set of ocular dominance columns, is also puzzling, but more experimental work is needed to clarify the nature of the difference.

Cats reared in the dark for four or more months are known to be abnormally sensitive to the effects of monocular deprivation, judged by changes in the responsiveness of cortical cells to visual stimuli⁸. This is at a time when similar deprivation would be without effect in a normal animal. The present results therefore support the idea that the critical period for monocular deprivation is linked to the time when left and right eye inputs to the cortex are incompletely segregated^{1,3,4}.

I thank Lyn Cummings for technical assistance and Ralph Freeman for helpful discussions. The research was supported by grants to Professor H. B. Barlow from the MRC and the Royal Society.

Received 10 December 1980; accepted 4 February 1981.

- Hubel, D. H., Wiesel, T. N. & LeVay, S. *Phil. Trans. R. Soc. B* **278**, 131–163 (1977).
- Shatz, C. J. & Stryker, M. P. *J. Physiol., Lond.* **281**, 267–283 (1978).
- LeVay, S., Stryker, M. P. & Shatz, C. J. *J. comp. Neurol.* **179**, 223–244 (1978).
- Swindale, N. V. *Proc. R. Soc. B* **208**, 243–264 (1980).
- LeVay, S., Wiesel, T. N. & Hubel, D. H. *J. comp. Neurol.* **191**, 1–51 (1980).
- Rogers, A. W. *Techniques of Autoradiography* 2nd edn (Elsevier, Amsterdam, 1973).
- Shatz, C. J., Lindström, S. & Wiesel, T. N. *Brain Res.* **131**, 103–116 (1977).
- Cynader, M. in *Developmental Neurobiology of Vision* Series A: Life Sciences Vol. 27 (ed. Freeman, R. D.), 109–120 (Plenum, New York, 1979).

Common precursor of lysozymes of hen egg-white and bacteriophage T4

B. W. Matthews, M. G. Grütter, W. F. Anderson* & S. J. Remington*

Institute of Molecular Biology and Department of Physics,
University of Oregon, Eugene, Oregon 97403, USA

The lysozymes of hen egg-white and bacteriophage T4 have similar catalytic properties but their amino acid sequences are not homologous. The question therefore arises whether they are derived from a common ancestral protein or have arisen independently. On the basis of the data we have gathered, it is shown here that the two enzymes are similar in the conformation of their backbones, in their modes of binding substrates, in specific protein-substrate interactions and in their presumed modes of action. We conclude that the two enzymes have diverged from a common precursor. This seems to be the most convincing example to date of the divergence of proteins with nonhomologous amino acid sequences.

If anything, the large number of protein structures and amino acid sequences now known has confused rather than simplified the main issue of whether extant proteins have evolved from a small number of ancestral precursors or have arisen independently. When the amino acid sequences are closely homologous, there is no difficulty: three-dimensional structures are always similar, and it is reasonable to infer that there is a common ancestral protein. In the absence of sequence homology, on the other hand, the origin of two proteins is logically ambiguous. Functional and/or structural similarity may indicate either divergent evolution from a common precursor or convergent evolution from amino acid sequences arising independently.

In many circumstances, tell-tale sequence homology is not readily apparent. The most notable example is the 'nucleotide-binding domain' of four NAD-dependent dehydrogenases and certain other enzymes¹⁻⁵. Although the amino acid sequences of lactate, alcohol and glyceraldehyde-3-phosphate dehydrogenases are not obviously homologous, when the sequences are aligned in such a way that common structural features coincide, a degree of sequence homology becomes evident, suggesting divergent evolution. In other cases of enzymes with similar structures but no apparent sequence homology, the question of divergent or convergent evolution has remained open⁶⁻²¹. The greater the structural correspondence between two proteins, as estimated by a variety of methods^{4,16,22-28}, the more likely are the two proteins to have evolved from the same precursor, but structural similarity *per se* is no proof of divergence. In some instances, two domains of a single protein have been observed to have structural similarity but not sequence homology. The fact that the polypeptide chains of the two domains are contiguous, even though their sequences are not related, has been taken as evidence for gene duplication, that is, a special case of divergent evolution^{9,28-30}.

In the case of bacteriophage T4 lysozyme (T4L) and hen egg-white lysozyme (HEWL), there are additional critical features common to the respective structures which strongly suggest that the two enzymes have diverged from a common precursor. Correspondence between HEWL and T4L can be seen at three levels: (1) the respective backbones, (2) the locations of saccharides bound in the respective active site clefts and (3) specific protein-substrate interactions and the probable mechanisms of catalysis.

Rossmann and Argos⁴ showed that, allowing for certain insertions and deletions, the amino-terminal half of T4L is structurally similar to HEWL. Of the 129 residues in HEWL and 164 in T4L, 78 α -carbon atoms were found to be 'equivalent' and could be superimposed within a r.m.s. distance of 4.1 Å. Notwithstanding the structural correspondence, the minimum base change per codon for the 78 'equivalent' residues is 1.53, a value expected for two amino acid sequences chosen at random. Remington and Matthews²⁵ used a different method of comparing the two lysozyme structures and concluded that the agreement between them was significantly better than one would expect by chance. The part of T4L 'equivalent' to HEWL is shown in Fig. 1.

Recently, we have completed a study of the binding of a series of mono- and oligosaccharides to phage lysozyme which has permitted a detailed comparison of the respective active sites of T4L and HEWL^{31,32}. We took the coordinates of the tetrasaccharide δ -lactone bound to HEWL³¹, and, using the respective transformations between HEWL and T4L proposed by Rossmann and Argos⁴ and Remington and Matthews²⁵, transformed these coordinates into the T4L coordinate frame. As summarized in Table 1, the centres of the four sugars agree within 1–2 Å with the saccharide binding sites determined experimentally for T4L. The experimental data provide striking confirmation of the sugar binding sites in T4L predicted by Rossmann and Argos, and support their conclusions concerning the relationship between the two lysozymes. Furthermore, there is remarkable agreement (Fig. 2) between parts of T4L and HEWL which are critically important for substrate binding and catalysis.

In the case of HEWL, Phillips and co-workers proposed a mechanism of action in which substrate strain in subsite D favours the formation of a carbonium ion intermediate which is stabilized by Asp 52, with Glu 35 acting as the proton donor^{33,34}. Subsequent work suggests that electronic stabilization rather than strain may be more important in promoting catalysis^{35,36}. In the case of T4L, we see a close parallel^{31,32}. Here, model building suggests that, as with HEWL, a saccharide with the normal chair conformation cannot be accommodated in site D because of steric interference, but a saccharide in the sofa conformation, as in the proposed transition state, will fit into the D site. For HEWL, the critical close contact in site D is between C₆₆ and the backbone C=O of Gln 57. As can be seen in Fig. 2, an exactly analogous close approach occurs for T4L between C₆₆ and the backbone C=O of Gly 30. In addition, for both enzymes the NH of the adjacent peptide forms a characteristic hydrogen bond to the acetamido group of saccharide C which is thought to help promote the preferential binding of monosaccharides in site C.

Finally, Asp 20 of T4L is in virtually the identical position, relative to substrate, as Asp 52, and is ideally placed to stabilize the developing positive charge on the postulated carbonium ion intermediate. Glu 11 of T4L is close to, but not in exactly the same relative position, as Glu 35 of HEWL, and is salt linked to Arg 45, in contrast to the nonpolar environment of Glu 35. Assuming a slight conformational adjustment when substrate is bound, Glu 11 could move into the required position to act as

Table 1 Comparison of predicted and observed saccharide binding sites for phage lysozyme

Saccharide subsite	Distance between predicted and observed site (Å)	
	Prediction based on 78 'equivalent' α -carbon atoms ^{4,32}	Prediction based on the best agreement between 80 contiguous α -carbon atoms ^{25,32}
A	1.8	2.6
B	1.5	1.6
C	1.5	1.1
D	1.5	1.2

* Present addresses: Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7 (W.F.A.); Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG (S.J.R.).

Fig. 1 Stereo view of the α -carbon backbone of bacteriophage T4 lysozyme showing the location of a bound oligosaccharide. The solid lines connect the 78 α -carbon atoms of phage lysozyme which have equivalent α -carbons in hen egg-white lysozyme⁴.

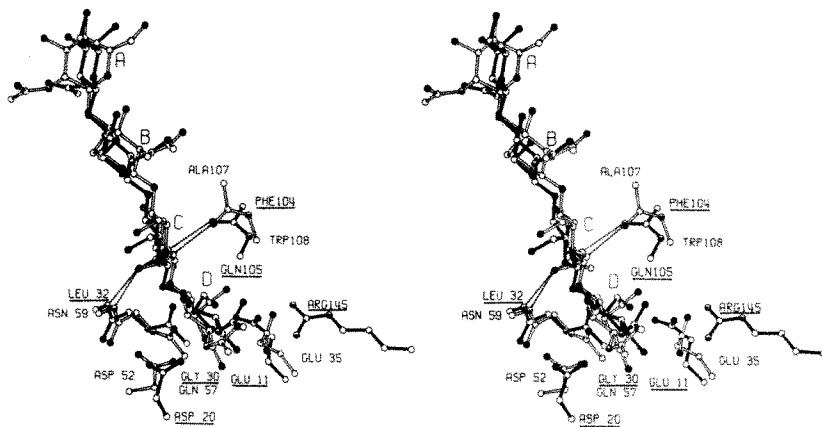


Fig. 2 Superposition of elements of the active site of hen egg-white lysozyme (open bonds), including a bound tetrasaccharide, on the corresponding elements of T4 phage lysozyme (solid bonds and names underlined).

the proton donor, as Glu 35 is postulated to do in HEWL. (Additional details are given in ref. 32.)

Thus, phage lysozyme and HEWL have similarities in their backbone conformations, in the locations of bound substrates and in the specific details of substrate binding and catalysis. These similarities strongly suggest that the two lysozymes have diverged from a common precursor. To argue for convergent, rather than divergent evolution, one would have to postulate that the similarities between the two lysozymes which we have enumerated all arose independently. We consider this unlikely.

The two lysozymes provide an example of the changes in proteins which can occur during evolution. In this case the only amino acids which seem to have retained their identities are Glu 11 (Glu 35) and Asp 20 (Asp 52). Other similarities between the respective lysozymes (Figs 1, 2) are at the level of the protein backbone rather than side-chain identity. As will be described in more detail elsewhere, the part of the upper lobe of T4L (Fig. 1), which has no counterpart in HEWL, seems to be involved in binding the cross-linking peptide of *Escherichia coli* cell walls. This peptide is required for substrates of T4L, but is not essential for HEWL.

After the submission of this manuscript, Jung *et al.*³⁷ reported the locations of the four exons of HEWL. Exon 2 codes for amino acids 28–32 and exon 3 for amino acids 82–108. Taken together, these two regions include the part of HEWL which is most similar to T4L and support the hypothesis³⁸ that exons may have been involved in the evolution of proteins. The same suggestion has been made independently by P. J. Artymiuk, C. F. Blake and A. E. Sippel (personal communication).

This work was supported in part by a Swiss NSF Fellowship to M.G.G., a NIH Postdoctoral Fellowship (GM 05972) to W.F.A., NIH research grants GM 20066 and GM 21967, NSF grant PCM-8014311 and a grant from the M. J. Murdock Charitable Trust. We thank Dr Colin Blake for communicating his manuscript before publication.

Received 15 September 1980; accepted 5 January 1981.

- Blake, C. C. F. *Nature* **250**, 284–285 (1974).
- Blake, C. C. F. *Nature* **267**, 482–483 (1977).
- Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* **250**, 194–199 (1974).
- Rossmann, M. G. & Argos, P. *J. molec. Biol.* **105**, 75–96 (1976).
- Rossmann, M. G., Liljas, A., Branden, C. I. & Banaszak, L. J. in *The Enzymes* Vol. 11, 3rd edn (ed. Boyer, P. E.) 61–102 (Academic, New York, 1975).
- Blake, C. C. F. & Evans, P. R. *J. molec. Biol.* **84**, 585–601 (1974).
- Banks, R. D. *et al. Nature* **279**, 773–777 (1979).
- Campbell, J. W., Watson, H. C. & Hodgson, G. I. *Nature* **250**, 301–303 (1974).
- Ploegman, J. H., Drent, G., Kalk, K. H. & Hol, W. G. J. *J. molec. Biol.* **123**, 557–594 (1978).
- Fletterick, R. J., Syguch, J., Semple, H. & Madsen, N. B. *J. biol. Chem.* **251**, 6142–6146 (1976).
- Weber, I. T. *et al. Nature* **274**, 433–437 (1978).
- Levine, M., Muirhead, H., Stammers, D. K. & Stuart, D. I. *Nature* **271**, 626–630 (1978).
- Eichele, G. *et al. J. molec. Biol.* **133**, 161–180 (1979).
- Borisov, V. V., Borisova, S. N., Sosfenov, N. I. & Vainshtein, B. K. *Nature* **284**, 189–190 (1980).
- Matthews, D. A. *et al. J. biol. Chem.* **253**, 6946–6954 (1978).
- Rossmann, M. G. & Argos, P. *J. molec. Biol.* **109**, 99–129 (1977).
- Rossmann, M. G. & Argos, P. *J. biol. Chem.* **250**, 7525–7532 (1975).
- Richardson, J. S., Richardson, D. C., Thomas, K. A., Silverton, E. W. & Davies, D. R. *J. molec. Biol.* **102**, 221–235 (1976).
- Phillips, D. C., Sternberg, M. J. E., Thornton, J. M. & Wilson, I. A. *J. molec. Biol.* **119**, 329–351 (1978).
- Argos, P., Rossmann, M. G. & Johnson, J. E. *Biochem. biophys. Res. Commun.* **75**, 83–86 (1977).
- Blow, D. M., Irwin, M. J. & Nyborg, J. *Biochem. biophys. Res. Commun.* **76**, 728–734 (1977).
- Schulz, G. E. & Schirmer, R. H. *Nature* **250**, 142–144 (1974).
- Richardson, J. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2619–2623 (1976).
- Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **105**, 357–382 (1976).
- Remington, S. J. & Matthews, B. W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2180–2184 (1978).
- Remington, S. J. & Matthews, B. W. *J. molec. Biol.* **140**, 77–99 (1980).
- Schulz, G. E. *J. molec. Biol.* **9**, 339–342 (1977).
- Schulz, G. E. *J. molec. Biol.* **138**, 335–347 (1980).
- McLachlan, A. D. *J. molec. Biol.* **128**, 49–79 (1979).
- McLachlan, A. D. in *Protein Folding* (ed. Jaenicke, R.) 79–99 (Elsevier, Amsterdam, 1980).
- Anderson, W. F., Grütter, M. G., Remington, S. J. & Matthews, B. W. *J. molec. Biol.* (in the press).
- Matthews, B. W., Remington, S. J., Grütter, M. G. & Anderson, W. F. *J. molec. Biol.* (in the press).
- Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C. & Tjian, R. *J. molec. Biol.* **88**, 349–371 (1974).
- Blake, C. C. F. *et al. Proc. R. Soc. B* **167**, 378–388 (1967).
- Warshel, A., Levitt, M. *J. molec. Biol.* **103**, 227–249 (1976).
- Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5250–5254 (1978).
- Jung, A., Sippel, A. E., Grez, M. & Schutz, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5759–5763 (1980).
- Gilbert, W. *Nature* **271**, 501 (1978).

Unusual filterable oncogenic agent isolated from horizontally transmitted Syrian hamster lymphomas

Joseph H. Coggin Jr*, John E. Oakes*,
Robert J. Huebner† & Ray Gilden‡

* Department of Microbiology and Immunology, University of South Alabama, College of Medicine, Mobile, Alabama 36688, USA

† National Cancer Institute, Building 37, Room 2D28, Bethesda, Maryland 20205, USA

‡ Frederick Cancer Center, Frederick, Maryland 21701, USA

An outbreak of horizontally transmitted malignant lymphoma in an experimental hamster holding facility was previously reported¹. Retroviridae (oncornavirus) or other conventional oncogenic viruses (oncoviruses) could not be detected in these lymphomas by immunological methods, direct isolation procedures or electron microscopy² but an infectious agent was clearly involved. The incidence of lymphomas during five recurrent epidemics ranged from 50 to >90% in young, inbred and random-bred Syrian golden hamsters exposed¹. The agent seemed to be resistant to UV inactivation, formaldehyde vapour and other viricidal agents (chlorine and iodine), and stable for long periods in the absence of hamster hosts in the contaminated facility^{1,2}. Associated disease syndromes in exposed hamsters included severe enteritis, pyelonephritis, the occasional appearance of warts, poor breeding efficiency and intussusception^{1,2}. We now report the successful, cell-free isolation of an unusual, filterable agent prepared in protamine sulphate buffer from primary and animal-passaged lymphomas, which produces lymphomas with good efficiency when injected subcutaneously (s.c.) into newborn Syrian inbred (LSH) and random-bred (LVG) hamsters. The agent could be reisolated from these induced lymphomas and injected into other hamsters to reproduce the neoplastic condition. It showed characteristics suggested for a mammalian viroid (a non-encapsidated, DNase-sensitive low-molecular-weight, disease-causing, self-replicating, naturally infectious nucleic acid).

Various isolation techniques³⁻⁶ were used in attempts to obtain cell-free, filterable agent responsible for lymphoma induction in young or newborn hamsters inoculated either s.c. or intraperitoneally (i.p.) (Table 1). None of these yielded cell-free oncogenic agent, but the incidence of 'background' lymphoma mediated by naturally transmitting agent among the hamsters housed in the contaminated facility was suppressed in hamsters challenged with non-oncogenic extracts of lymphoma, compared with placebo controls which received buffer or normal spleen extracts. In this report, data are given for the incidence of background lymphomas detected among control or test hamsters placed in the agent-contaminated facility. The incidence of background lymphomas can range from 0 to >50% in some rooms. Such spontaneous lymphomas contracted from agent in the facility were unavoidable because we could not use new facilities for each experiment. Previous work demonstrated that it was difficult to rid the unit of contaminating agent by extensive sanitization that would destroy most viruses. However, background lymphomas were readily discerned from protamine sulphate (PS)-stabilized agent induced tumours by site of appearance.

We obtained cell-free agent using a simple extraction method in which PS (0.01 mg ml⁻¹) was added to the extraction buffer (Fig. 1). Filtered (0.45–0.1 µm) extracts of lymphoma prepared with PS were able to produce 40–95% lymphomas in trials carried out over 2 yr when inoculated s.c. into newborn hamsters. Two additional PS preparations were tested s.c. in newborns in a new facility which was free of contamination and

Table 1 Protection of hamsters against contagious lymphoma after injection of Moloney extracts of spontaneous intestinal lymphoma

Preparation	Hamsters with lymphomas/ Test extract (%)	Control buffer solution (%)
Moloney extract:		
Trial 1	4/18 (22)	9/9 (100)
Trial 2	0/16 (0)	4/11 (36)
Trial 3	2/34 (6)	3/9 (33)
Freeze-thaw extract	3/34 (9)	4/9 (44)
Saline homogenate	1/11 (10)	3/9 (33)
Phenol extraction*		
Trial 1	0/20	0/4
Trial 2	0/30	0/7
Trial 3	0/37	NT
Hirt extraction*		
Trial 1	0/10	2/10 (20)
Trial 2	0/23	1/9 (11)

The Moloney extract comprised 15 g of lymphoma extracted in citrate buffer and hyaluronidase (for methods see Fig. 1 legend). 0.5 ml was injected into the right rear legs of 3–4-week-old LVG hamsters. All solutions were passed through a 0.45-µm filter. All lymphomas recorded were at node sites in the peritoneal cavity or in the cervical-thymic node area. The freeze-thaw extract comprised 15 g of lymphoma extracted after three repetitive flash freezes (–80 °C) and slow thaw (25 °C) in phosphate-buffered saline (PBS), pH 6.8. The solution was passed through a 0.45-µm filter and 0.05 ml injected i.p. The saline homogenate comprised 15 g of lymphoma extracted by homogenization in PBS and filtered through a 0.45-µm membrane. 0.05 ml was injected i.p. Each group originally comprised 40–48 newborns. The survivors were rapidly reduced to those given by enteritis. Young hamsters which died from enteritis had no lymphomas at the time of death (2–6 weeks). Enteritis was not seen after 5–6 weeks, and was observed equally in control groups and in hamsters injected with lymphoma extracts. These trials were carried out in a new trailer unit. NT, not tested.

*Lymphomas in these groups were detected in the lower flank, adjacent to the injection site.

were oncogenic. These tumour preparations were from cell-free, PS agent induced lymphomas. When the same PS-stabilized extracts were injected i.p., there were fewer tumours in the injection area than after s.c. injection (data not shown).

Electron microscopy using several staining methods^{7,8} on spontaneous lymphomas, animal-passaged lymphomas and concentrated PS extracts of lymphoma oncogenic in newborn hamsters were devoid of virions unique to the lymphomas. R or H particles and rare budding particles were infrequently observed in normal and fibroblastic tissue and epithelial cells from lymphomatous hamsters, but these were also commonly detected in non-exposed hamsters⁹ from breeder sources. We concluded that there were no virions or virus-like particles (VLPs) unique to lymphoma tissue or to uninvolved hair follicles, skin, spleen, gut, or normal lymphocytes of lymphoma-bearing hamsters.

Preparations of lymphoma or serum from lymphoma-bearing hamsters were tested for mycoplasma, antibody to lymphocytic choriomeningitis virus, SV40, polyoma and SV5 and Reo virus, with negative results. Antibody to Sendai virus was present in healthy, breeder-supplied hamsters. In repeated trials, serum from either lymphoma-bearing or their exposed, disease-free cage-mates failed consistently to neutralize PS-stabilized agent and prevent lymphoma induction.

Hamster age was most important in maximum sensitivity to s.c. oncogenic challenge with PS-stabilized, filtered extracts of lymphoma (Table 2), with 1-day-old, but not 10-day-old, hamsters being most susceptible. The oncogenic agent in PS was readily inactivated by DNase but not by RNase or proteases (Table 3).

Diener¹⁰ concluded that a viroid etiology should be considered when: (1) no microorganism was consistently associated with the disease; (2) VLPs could not be identified in extracts or sections of diseased tissue by electron microscopy; (3) the

infectious agent was difficult to pellet by ultracentrifugation; and (4) nuclease treatment inactivated the agent. Fulfilment of these criteria alone does not constitute absolute proof of a viroid aetiology which depends on demonstrating non-encapsidated nucleic acid¹⁰.

We suggest that this lymphoma agent is a naturally transmitting, oncogenic, naked DNA viroid-like agent whose origin and mode of transmission remain unclear. The agent is difficult to sediment in buffer, DNase labile, produces no unique virus particles in lymphoma tissue or normal tissues of the tumour-bearing host nor in concentrated, filtered oncogenic extracts, and requires PS for cell-free stabilization. It seems to be unrelated to known oncogenic Retroviridae or papovaviruses, and generates no neutralizing antibody in trials using tumour-bearer or disease-free cage-mate serum. We must still prove the existence of infectious nucleic acid to claim a viroid status for the oncogenic agent.

Graffi *et al.*¹¹⁻¹³ reported a similar disease in Syrian hamsters and attributed the cause to an unidentified budding virus of the Retroviridae family. They suggested that various agents, such as

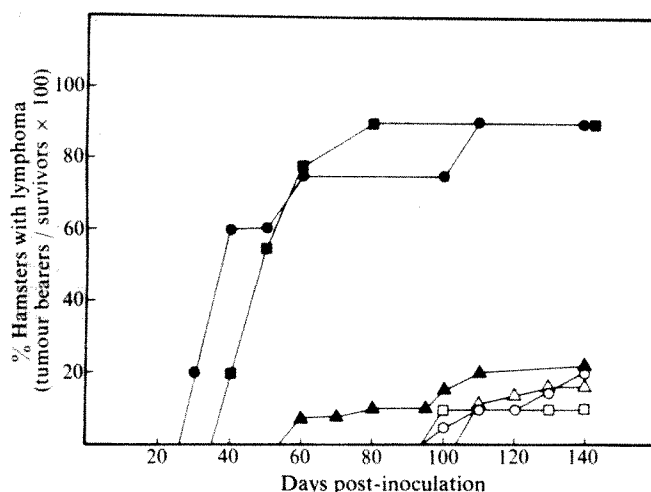


Fig. 1 Lymphoma production using extracts of hamster lymphoma. ○, Control (no inoculations). Hamsters inoculated s.c. in rear flank: ▲, Moloney method; ■, PS method (LSH recipients); ●, PS method (LVG recipients); △, control (Moloney extraction buffer only); □, control (PS solution only). Modified Moloney extraction method: a 6.6% suspension was prepared of intestinal lymphoma which had been finely minced in 0.153 M potassium citrate containing 1 mg of hyaluronidase per 100 ml. This was digested for 1 h at 25 °C with occasional mixing in a blender (15-s pulse) and homogenized in the blender in an ice bath (3 min). The extract was centrifuged at 2,500 r.p.m. for 20 min at 4 °C and the middle supernatant (S_1) collected. This was centrifuged again as above and the middle supernatant (S_2) collected. S_2 was centrifuged at 17,000g for 1 min at 4 °C and the supernatant (S_3) collected. S_3 was filtered through a 8- μ m Biorad filter, then through a 1- μ m Biorad filter, and finally a 0.45- μ m Millipore filter. The S_3 filtrate was centrifuged at 40,000g for 1 h at 4 °C and the pellet rinsed three times with 0.05 M sodium citrate buffer, pH 6.7. The pellet was resuspended in citrate buffer using 1 ml per g of original lymphoma and exposed to three 1-s bursts with a sonicator to achieve cavitation. 0.1 ml was injected s.c. Protamine sulphate method: lymphoma (10–15 g) was minced to a fine suspension in Hank's balanced salt solution, and the final suspension forced through an 18-gauge needle. The cells were pelleted at 1,500 r.p.m. for 10 min at 4 °C and the pellet washed twice with 5 ml PBS, pH 6.8, per 5 g tumour used. The pellet was resuspended in PBS containing 0.01 mg ml⁻¹ PS, 0.01 M Tris, 0.001 M EDTA (pH 8.0) and kept at 4 °C for 30 min. This was homogenized in a Dounce unit and centrifuged at 2,000 r.p.m. for 10 min at 4 °C. The supernatant (S_1) was centrifuged in a 50 Ti Beckman rotor at 10,000g ($\omega^2 t = 365$) for 60 min at 4 °C. The supernatant (S_2) was filtered through a 0.45- μ m Millipore filter, then a 0.22- μ m filter. 0.1 ml was injected s.c.

Table 2 Relationship of age to susceptibility to lymphoma induction by cell-free, filtered preparations from spontaneous lymphoma

Protamine extract of:	Age of challenge (days post-birth)	Final tumour incidence (tumours/survivors, %)
Normal hamster spleen	1	2/17 (12)
	5	1/12 (8)
	10	3/22 (9)
Buffer alone	1	3/23 (13)
	5	0/11 (0)
	10	2/20 (10)
Spontaneous lymphoma	1	14/20 (70) $P < 0.01^*$
	5	9/22 (41) $P < 0.05$
	10	5/24 (21) $P > 0.05$ NS

A spontaneous lymphoma which appeared in the lymph node of the neck area (weight 5.3 g) was removed and extracted using the PS extraction method (see Fig. 1 legend). 0.05 ml of the final supernatant was injected s.c. into the lower rear quadrant of hamsters. Hamsters were challenged with extracts on the day indicated. All mothers were obtained from the Charles River Hamstery as primiparous and all groups were weaned and sex segregated at 21 days of age. Hamsters were palpated and examined weekly for evidence of tumour. The experiments were terminated at 222 days after birth and survivors autopsied for evidence of tumours. The last tumour was detected at 178 days and the first appeared at 28 days. NS, not significant.

* Significance by χ^2 compared with normal spleen cell extract.

Table 3 Susceptibility of oncogenic agent-producing lymphomas to DNase

Treatment	Cumulative no. lymphomas/ no. survivors Days post-injection of neonate LVG hamsters			
	0	80	100	150
None—filtered PS-stabilized lymphoma extract alone (30 min, 37 °C)	0/17	4/9	7/9	8/9
DNase (30 min, 37 °C) plus filtered PS-stabilized lymphoma extract	0/20	0/16	0/16	0/15
Trypsin (0.25%; 20 min, 37 °C) plus filtered PS-stabilized lymphoma extract	0/16	8/16	16/16	—
RNase (30 min, 37 °C) plus filtered PS-stabilized lymphoma extract	0/10	3/9	4/6	6/6
Heat-inactivated DNase* (30 min, 37 °C) plus filtered PS-stabilized lymphoma extract	0/13	3/12	8/9	9/9
Control—PS-containing extraction solution only (no tumour)	0/15	0/14	1/14	1/14

DNase, deoxyribonuclease 1, type 1, bovine pancreas (Sigma); RNase free, 250 μ g ml⁻¹. As little as 50 μ g ml⁻¹ gave the same result in another trial. Trypsin, type IX (Sigma), 0.5%. Papain was also ineffective in another trial. RNase, ribonuclease, type A-1, protease free (Sigma).

* DNase was inactivated by boiling for 10 min before use.

human tumour DNA and hamster papilloma DNA, could activate an endogenous C-type virus which induced the lymphomas observed. We find no evidence for such a retrovirus in our epidemic disease. M. Gardner (personal communication) found that saline could activate lymphoma in Graffi's hamster stock; therefore it is possible that an agent identical to the one described here was being transmitted unrecognized in the Graffi hamster.

There are several problems in resolving the nucleic acid nature of this lymphoma-inducing agent: (1) it is highly stable in the environment (facility contamination); (2) the bioassay used

to detect the agent is long-term and expensive; (3) agent-producing cell cultures of the lymphoma have not been obtained despite extensive culture efforts—cells fail to grow *in vitro*; (4) reliable *in vitro* transformation of target cells by the agent has not been detected; (5) antigenic markers for the agent on transformed lymphoma cells are not yet sufficiently developed; and (6) other potential mammalian hosts are not susceptible to the agent. Work is being done to characterize this infectious, oncogenic DNA agent.

We thank Dr T. O. Diener for helpful comments during preparation of the manuscript. This work was supported by NCI grant CA 23491, NIH.

Received 5 September 1980; accepted 2 January 1981.

1. Ambrose, K. R. & Coggin, J. H. *J. natn. Cancer Inst.* **54**, 877–880 (1975).
2. Coggin, J. H., Thomas, K. V. & Huebner, R. *Fedn Proc.* **70**, 2086–2088 (1978).
3. Moloney, J. B. *J. natn. Cancer Inst.* **24**, 933–937 (1960).
4. Moloney, J. B. *Fedn Proc.* **21**, 19–21 (1962).
5. Green, M., Wold, W. S. M., Mackey, J. K. & Rigden, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6606–6610 (1979).
6. Hirt, B. *J. molec. Biol.* **26**, 365–369 (1967).
7. Fresca, J. M. & Parks, V. R. *J. Cell Biol.* **25**, 157–163 (1965).
8. Huxley, H. E. & Zubay, G. *J. biophys. biochem. Cytol.* **11**, 273–278 (1961).
9. Bernhard, W. & Tournier, P. *Annls Inst. Pasteur Paris* **107**, 447–452 (1964).
10. Diener, T. O. *Viroids and Viroid Diseases* 1–224 (Wiley, New York, 1979).
11. Graffi, A., Bender, E., Schramm, T., Kuhn, W. & Schneiders, F. *Proc. natn. Acad. Sci. U.S.A.* **64**, 1172–1175 (1969).
12. Graffi, A. *et al. Br. J. Cancer* **22**, 577–581 (1968).
13. Graffi, A., Schramm, T., Graffi, I., Bierwolf, D. & Bender, E. *J. natn. Cancer Inst.* **40**, 867–873 (1968).

$(\text{Na}^+ + \text{K}^+)\text{ATPase}$ has one functioning phosphorylation site per α subunit

W. H. M. Peters, H. G. P. Swarts,
J. J. H. M. de Pont,
F. M. A. H. Schuurmans Stekhoven
& S. L. Bonting

Department of Biochemistry, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands

$(\text{Na}^+ + \text{K}^+)\text{ATPase}$ contains two different subunits, a catalytic subunit (α) and a subunit with uncertain function (β). The enzyme binds ATP, ouabain and vanadate, and can be phosphorylated by ATP as well as by inorganic phosphate. From the previously reported maximal binding and phosphorylation capacities of 3.5–4.3 nmol P per mg protein^{1–3} (based on Lowry protein determination) and the earlier molecular weight value of ~250,000^{1,2}, a molar binding and phosphorylation capacity of 0.87–1.07 mol per mol enzyme was derived. As it is generally agreed that the enzyme molecule contains two α subunits^{1–6} or even a multiple of two^{7,8}, it has been suggested that the enzyme operates by means of a so-called 'half-of-the-sites mechanism', whereby only one of the two α subunits can be phosphorylated at any one time^{9,10}. We now present evidence that every α subunit can be phosphorylated simultaneously, which rules out the operation of such a mechanism.

We have recently shown that the molar ratio of the α and β subunits in the enzyme from the rabbit kidney is 1:1 (ref. 11). We have also determined the molecular weights of the subunits by sedimentation equilibrium analysis in the absence of detergent. After correction for carbohydrate and residual phospholipid contents, we obtained protein molecular weights of 120,600 for the α subunit and 42,800 for the β subunit. This means that the smallest unit, containing the catalytic subunit,

would be an $\alpha\beta$ dimer with a protein molecular weight of 163,400.

During these studies we noticed considerable discrepancy between results of the commonly used modified Lowry method¹² with bovine serum albumin as a standard and the absolute method by means of quantitative amino acid analysis. Recently, Craig and Kyte⁶ have also indicated that the Lowry method may give falsely high results for this type of membrane protein. We have, therefore, compared the Lowry protein values for a number of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ preparations with those from the absolute method of quantitative amino acid analysis in the presence of an internal standard (Table 1). The protein values obtained with the Lowry method are 1.345 (s.e.m. 0.018, $n = 6$) times greater than those obtained with the absolute method.

This finding has far-reaching consequences for several parameters of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$, such as its turnover number and specific activity, and especially its phosphorylation capacity. Nearly all published values for these parameters are based on Lowry protein determinations according to the original procedure¹³ or to modifications of this method; it would now seem that these values are generally too low.

We have also reinvestigated the phosphorylation capacity and found that the removal of endogenous ATP, which is added during purification of the enzyme to stabilize it, lowers the specific activity by 21% (s.e.m. 6, $n = 11$). We have, therefore, determined the phosphorylation capacity both with and without prior removal of endogenous ATP. Results, plotted as phosphorylation capacity against specific activity for a given preparation, are shown in Fig. 1, with technical details in the legend. A linear relationship between these two parameters is found, with the values obtained with or without removal of ATP before phosphorylation falling along the same line. For the most active preparations a phosphorylation capacity of 5.6 nmol P per mg protein (based on amino acid analysis) is obtained, slightly

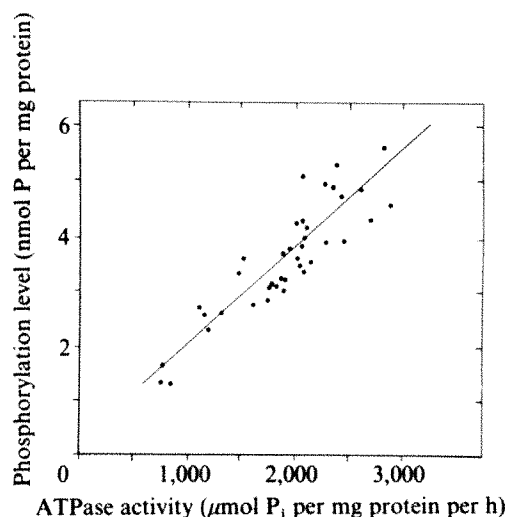


Fig. 1 Relationship between phosphorylation capacity and specific activity of 41 purified $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ preparations from rabbit kidney. The enzyme was prepared as previously described¹². Activity was determined according to Schoot *et al.*¹⁵ and protein by quantitative amino acid analysis as described in Table 1 legend. Phosphorylation was carried out as described before¹⁶ (34 preparations) or with the modification that the enzyme is phosphorylated without prior removal of endogenous ATP (seven preparations). In the latter case excess $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (1–2.5 mM, diluted with cold ATP) was added. In both cases there is the same linear relationship between phosphorylation level and specific activity of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$. The equation for the regression line, calculated by the method of least squares, is $y = 1.737 \cdot 10^{-3} x + 0.219$ ($r = 0.90$).

Table 1 Protein content of rabbit kidney ($\text{Na}^+ + \text{K}^+$)ATPase determined by the Lowry method and by amino acid analysis

Preparation no.	Lowry method (mg protein ml ⁻¹)	Amino acid analysis (a.a.a.)	Ratio Lowry/a.a.a.
1	4.80	3.67	1.308
2	1.52	1.14	1.333
3	2.01	1.47	1.367
4	2.03	1.48	1.372
5	1.67	1.19	1.403
6	2.05	1.59	1.289
Mean			1.345 ± 0.018 (s.e.m.)

A 100- μl aliquot of a ($\text{Na}^+ + \text{K}^+$)ATPase suspension (~ 2 mg protein ml⁻¹) in 0.25 M sucrose, 10 mM Tris-HCl, 1 mM EDTA (pH 7.4) was added to a glass tube containing 150 nmol norleucine, followed by 300 μl 7.6 M HCl. The tube was vacuum sealed and hydrolysed in the dark for 24 h at 105 °C. After 10 min centrifugation at 20,000g, amino acid analysis of the supernatant was carried out in a modified JEOL type JLC-6AH amino acid analyser. The results of the amino acid analysis were corrected for incomplete hydrolysis of valine and isoleucine and degradation of serine and tryptophan¹¹. The protein content was calculated by summation of the amino acid weights. Parallel samples of the ($\text{Na}^+ + \text{K}^+$)ATPase suspension were analysed for protein according to Lowry *et al.*¹³ after trichloroacetic acid precipitation as described by Jørgensen¹².

higher than the recently reported¹⁴ value of 5.3 nmol P per mg protein.

The maximal phosphorylation capacity of 5.6 nmol P per mg protein at an $\alpha\beta$ dimer molecular weight of 163,400 corresponds to 0.92 phosphorylation site per $\alpha\beta$ dimer. As even a highly purified enzyme preparation is not 100% pure on a protein basis (as seen by the appearance of minor protein bands beyond the two prominent bands belonging to the α and β subunits on SDS-gel electrophoresis, data not shown), this value is a slight underestimation of the true phosphorylation capacity.

Thus we conclude that every $\alpha\beta$ unit has one active phosphorylation site. As the molar binding values for ATP, ouabain and vanadate equal the phosphorylation level¹⁻⁷, our results also indicate that each $\alpha\beta$ unit can bind one molecule each of ATP, ouabain and vanadate. As the most likely subunit composition of the enzyme is $\alpha_2\beta_2$ (ref. 11), we may safely conclude that the $\alpha_2\beta_2$ tetramer has two phosphorylation sites, which can be phosphorylated simultaneously. This rules out the operation of a 'half-of-the-sites mechanism' in ($\text{Na}^+ + \text{K}^+$)ATPase from rabbit kidney.

We thank Annie M. M. Fleuren-Jakobs for technical assistance.

Received 4 November 1980; accepted 9 February 1981.

- Jørgensen, P. L. *Physiol. Rev.* **60**, 864-917 (1980).
- Peterson, G. L., Ewing, R. D., Hootman, S. R. & Conte, F. P. *J. biol. Chem.* **253**, 4762-4770 (1978).
- Perrone, J. R., Hackney, J. F., Dixon, J. F. & Hokin, L. E. *J. biol. Chem.* **250**, 4178-4184 (1975).
- Esmann, M., Skou, J. C. & Christiansen, C. *Biochim. biophys. Acta* **567**, 410-420 (1979).
- Hastings, D. F. & Reynolds, J. A. *Biochemistry* **18**, 817-821 (1979).
- Craig, W. S. & Kyte, J. J. *J. biol. Chem.* **255**, 6262-6269 (1980).
- Hansen, O., Jensen, J., Nørby, J. G. & Ottolenghi, P. *Nature* **280**, 410-412 (1979).
- Askari, A. & Huang, W. *Biochem. biophys. Res. Commun.* **93**, 448-453 (1980).
- Stein, W. D., Lieb, W. R., Karlisch, S. J. D. & Eilan, Y. *Proc. natn. Acad. Sci. U.S.A.* **70**, 275-278 (1973).
- Repke, K. R. H. & Schön, R. *Acta biol. med. germ.* **31**, K19-K30 (1973).
- Peters, W. H. M., De Pont, J. J. H. H. M., Koppers, A. & Bonting, S. L. *Biochim. biophys. Acta* **641**, 55-71 (1981).
- Jørgensen, P. L. *Biochim. biophys. Acta* **356**, 36-52 (1974).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
- Askari, A., Huang, W. & Anteau, J. M. *Biochemistry* **19**, 1132-1140 (1980).
- Schoot, B. M., Schoots, A. F. M., De Pont, J. J. H. H. M., Schuurmans Stekhoven, F. M. A. H. & Bonting, S. L. *Biochim. biophys. Acta* **483**, 181-192 (1977).
- Schuurmans Stekhoven, F. M. A. H., Swarts, H. G. P., De Pont, J. J. H. H. M. & Bonting, S. L. *Biochim. biophys. Acta* **597**, 100-111 (1980).

Diversity of immunoglobulin expression in leukaemic cells resembling B-lymphocyte precursors

L. B. Vogler, J. L. Preud'homme & M. Seligmann

Laboratory of Immunochemistry and Immunopathology (INSERM U108), Institut de Recherches sur les Maladies du Sang de l'Université Paris VII; Laboratoire d'Oncologie et d'Immuno-Hématologie du CNRS, Hôpital Saint Louis, 75475 Paris Cedex 10, France

W. E. Gathings, W. M. Crist & M. D. Cooper

Cellular Immunobiology Unit, Department of Pediatrics and Microbiology and the Comprehensive Cancer Center, University of Alabama in Birmingham, Birmingham, Alabama 35294, USA

F. J. Bollum

Department of Biochemistry, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20014, USA

Approximately 20% of patients with acute lymphocytic leukaemia (ALL) have leukaemic blasts with features of pre-B cells¹ which are the recently characterized precursors of B lymphocytes in normal development (for a review, see ref. 2). Pre-B cells isolated from normal bone marrow or fetal liver, and malignant cells from patients with pre-B cell leukaemia, are rapidly dividing lymphoid cells that contain cytoplasmic immunoglobulin μ heavy chains, but have no detectable surface immunoglobulin. The resemblance of immunoglobulin-containing ALL cells to normal precursors of B lymphocytes and their availability in relatively pure preparations allowed us to explore them as models of early stages in the differentiation of the B-lymphocyte line. We report here observations on the occurrence of intermediate pre-B/B-cell phenotypes, immunoglobulin isotype switching and the asynchrony of immunoglobulin heavy and light chain expression in 30 cases of ALL and 3 cases of chronic myelogenous leukaemia in lymphoblastic crisis (CML-BC).

Heparinized bone marrow from patients with untreated ALL or with CML-BC was examined after lysis of erythrocytes with a hypotonic solution of ammonium chloride. Surface immunoglobulins were detected by immunofluorescent staining of fresh, viable cells with fluorochrome-conjugated, affinity-purified goat anti-human μ , γ , α , κ and λ antibodies, as described elsewhere³. Intracytoplasmic immunoglobulin was detected by immunofluorescent staining of cells that had been centrifuged or smeared on to glass slides, fixed in cold 95% ethanol, 5% acetic acid and washed in buffered saline as described elsewhere¹. Differently coloured fluorochromes (rhodamine and fluorescein isothiocyanates) were used to stain cells before and after fixation, which allowed us to distinguish between surface and cytoplasmic determinants in single cells. In some patients observations were reproduced using a completely different set of reagents: fluorochrome-conjugated F(ab')₂ fragments of rabbit anti-human μ , γ , α , κ and λ antibodies as described elsewhere⁴. Terminal deoxynucleotidyl transferase was detected by indirect immunofluorescence using a rabbit anti-terminal deoxynucleotidyl transferase serum⁵ and fluorescein-conjugated goat anti-rabbit Fab. Results of immunofluorescence studies represent concurrent observations of at least two of the authors.

In 1 year we found 33 patients with intracytoplasmic immunoglobulin-positive lymphoblastic leukaemia. In 31 of these at least 10% of marrow blasts contained cytoplasmic and no surface immunoglobulin, as originally observed in pre-B

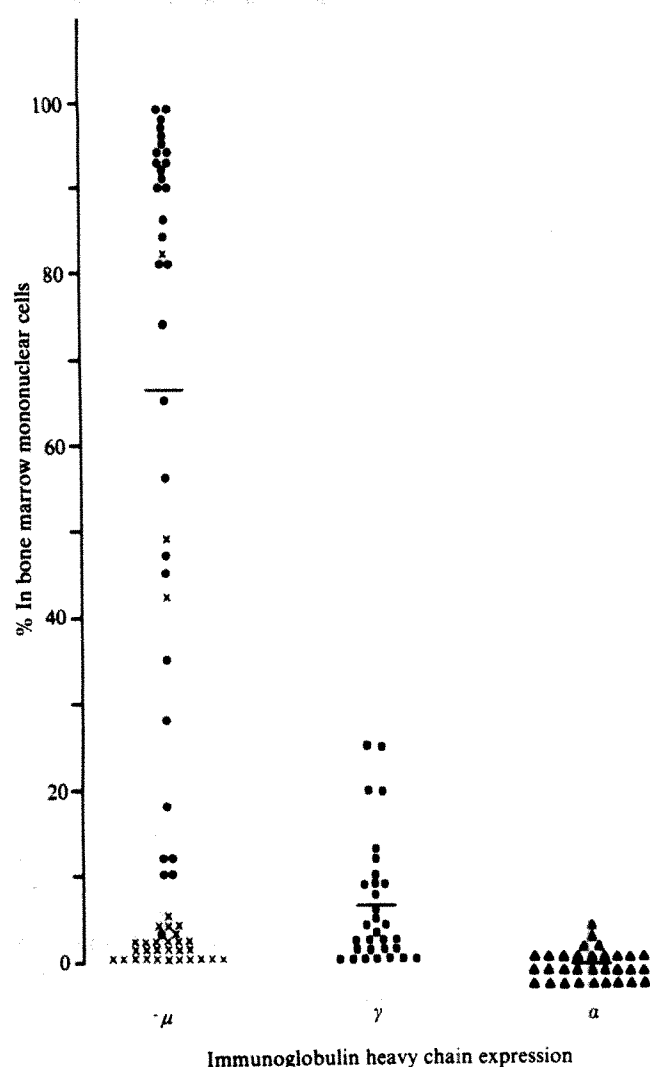


Fig. 1 Surface and cytoplasmic expression of immunoglobulin isotypes in leukaemic cells. ×, Surface μ ; ●, ■, ▲, cytoplasmic μ , γ and α chains, respectively. Horizontal bars represent the mean percentage of marrow mononuclear cells expressing intracytoplasmic immunoglobulin of each class.

leukaemia. In contrast, normal pre-B cells comprise only $1.0 \pm 0.6\%$ of human marrow mononuclear cells⁶. When taken collectively, blasts of pre-B phenotype accounted for 66.7% of the leukaemic cells in these patients (Fig. 1). In individuals, however, leukaemic cells of identical morphology were sometimes found to be composed of three distinct populations based on immunoglobulin expression: cells showing no immunoglobulin, cells with intracytoplasmic immunoglobulin only, and cells with both intracytoplasmic and scant amounts of surface μ . In three patients cells of the latter type accounted for a large proportion (42, 49 and 82%) of the lymphoblasts (Fig. 1). This phenotype, also observed by Brouet *et al.*⁷ and Greaves *et al.*⁸, is thought to represent cells at a stage of differentiation intermediate between that of the pre-B and B-cell stages. Several features distinguish these 'intermediate' cells from mature B lymphocytes: they have persistent expression of cytoplasmic μ chains and the nuclear enzyme, terminal deoxynucleotidyl transferase; and in two of three cases the sparse amounts of μ chain on the cell surface were not associated with detectable light chains.

As in previous studies^{1,7}, many patients with pre-B leukaemia had a very small proportion of a leukaemic cells that expressed

cytoplasmic γ or α chains in addition to cytoplasmic μ . However, in five of our present patients, cells containing γ chains accounted for a relatively large proportion ($>10\%$) of the leukaemic population, sometimes outnumbering μ -containing leukaemic cells (Fig. 2, Table 1).

In addition to the 31 patients already described, one with ALL and one with CML-BC had a small proportion of lymphoblasts (8% and 6%, respectively) that contained only γ heavy chains. Unlike the previous cases, μ chains (and other non- γ chains) were undetectable in both the γ -positive and γ -negative lymphoblasts from these individuals. As γ -containing cells comprised $<10\%$ of mononuclear cells in the marrow of these patients and μ chains were not detected, they did not conform to our original definition of pre-B cell leukaemia. However, there is evidence for the immaturity of these leukaemic cells and for their distinction from IgG-secreting lymphoplasmablasts: (1) the distribution and intensity of immunofluorescent staining with a pattern identical to that of μ -staining in pre-B cells; (2) the co-expression of cytoplasmic γ and nuclear terminal deoxynucleotidyl transferase in double immunofluorescence studies; and (3) the absence of detectable light chains in freshly collected lymphoblasts. Note also that κ chains were found in 10% of the γ -positive lymphoblasts from the patient with CML-BC when leukaemic cells were cultured for 24 h in RPMI 1640 medium supplemented with fetal calf serum.

Cells from 15 patients were examined for surface and cytoplasmic κ and λ light chains (Table 1). In only four of these were cytoplasmic light chains found in most of leukaemic cells containing μ heavy chains. Light chains usually were present in only a small proportion of cells containing heavy chains, or were undetectable. In five patients (see Table 1, patients 1, 2, 4, 10 and 11) when leukaemic cells were examined simultaneously for light and heavy chains by double immunofluorescent staining, the expression of cytoplasmic κ chains was associated with that of γ heavy chains. No surface light chains were detected in the absence of surface heavy chains.

Analysis of leukaemic cells from our patients with ALL and CML-BC reveals that immunoglobulin expression in the pre-B group of leukaemias is less homogeneous than originally thought (Fig. 2). Cells carrying sparse surface μ , probably representing an intermediate pre-B/B-cell stage of development, were predominant in three patients. The infrequent expression of light chains relative to μ chains in pre-B leukaemic cells from most of these patients agrees with similar results reported by Brouet *et al.*⁷ and Greaves *et al.*⁹. An asynchrony in the expression of cytoplasmic μ heavy chains and light chains has also recently been observed in murine cells of pre-B phenotype from normal fetal liver¹⁰, Abelson virus-infected bone

Table 1 Cytoplasmic immunoglobulin expression by lymphoblasts from patients with ALL or CML-BC

Patient	Diagnosis	Ig heavy chain			Ig light chain	
		μ	γ	α	κ	λ
1	ALL	10	25	<1	25	<1
2	ALL	81	10	<1	15	<1
3	ALL	92	<1	<1	5	<1
4	ALL	36	13	<1	14	<1
5	ALL	99	5	<1	<1	<1
6	ALL	95	3	<1	<1	<1
7	ALL	93	4	2	3	3
8	ALL	90	<1	<1	70	<1
9	ALL	86	4	<1	87	<1
10	ALL	65	25	<1	25	<1
11	ALL	10	20	<1	20	<1
12	CML-BC	81	<1	<1	<1	<1
13	CML-BC	3	9	<1	2	<1
14	ALL	<1	8	1	2	<1
15	CML-BC	<1	6	<1	1	<1

Results are expressed as percentages of positive cells among bone marrow mononuclear cells.

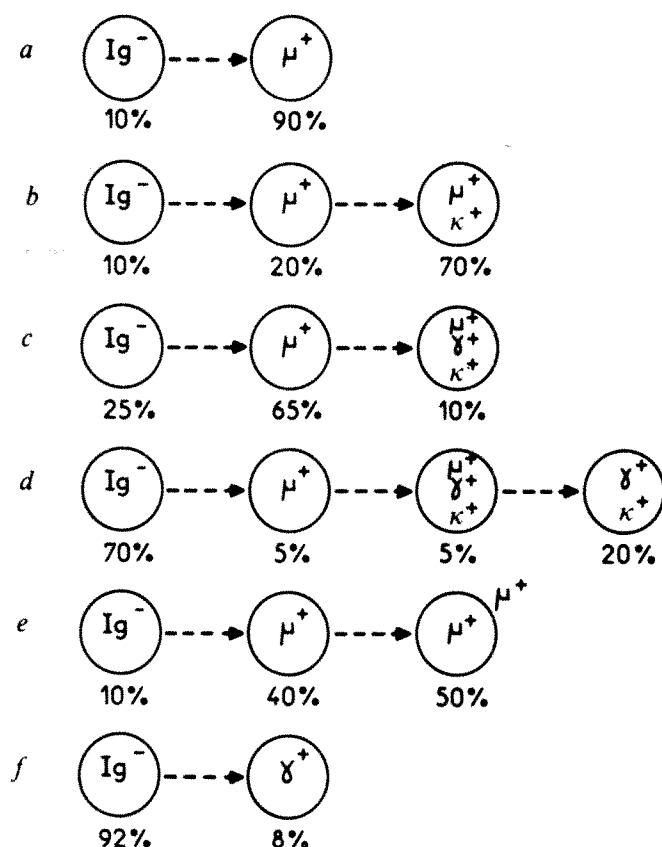


Fig. 2 Phenotypes observed in immunoglobulin-containing lymphoblastic leukaemias. Six patients were selected to demonstrate the heterogeneity of immunoglobulin expression in immunoglobulin-positive lymphoblastic leukaemias. A representative cell phenotype comprises at least 5% of the leukaemic cell population. In each case a varied number of leukaemic blasts expressed no detectable immunoglobulin and could be described as 'null' phenotype. The surface μ⁺ leukaemic cells in patient *e* may represent a developmental stage intermediate between that of pre-B and B cell. Both these cells and the γ⁺ cells in patient *f* had no detectable light chains, but contained nuclear terminal deoxynucleotidyl transferase. Dashed arrows show suggested directions of differentiation.

marrow-derived cell lines¹¹, and pre-B/myeloma cell hybridomas¹². In humans the study of bone marrow and fetal liver also indicates that only a small minority (<10%) of normal pre-B cells co-express κ or λ light chains with μ heavy chains; most normal human pre-B cells express only μ (H. Kubagawa, unpublished results). In contrast to these results and observations in most patients with pre-B leukaemia, κ light chains were detected in >5% of leukaemic cells from seven of our patients, usually in μ-containing cells, and their incidence coincided with the expression of γ chains in five patients. However, the absence of γ chains in cytoplasmic μ⁺, κ⁺ cells from two patients, and the absence of light chains in γ⁺ cells from two others, precludes any simple explanations for γ and light-chain concurrence during differentiation.

The expression of a second heavy chain isotype in some of our patients' leukaemic cells has not been previously observed in normal human pre-B cells. There are two possible explanations for this: because γ chains were observed in a small population of leukaemic cells from a minority of patients with pre-B leukaemia, class switching may be a rare event at the pre-B stage of differentiation. Therefore, it would be difficult to detect a small population of γ-containing cells among the ~1% of μ-positive pre-B cells present in normal marrow using conventional tech-

niques, although we have made considerable efforts in this direction. Alternatively, the γ chains found in these leukaemic cells may reflect abnormal regulation of immunoglobulin gene expression related to malignant transformation, rather than the sequential events of normal B-lymphocyte development.

Whichever the case, our results indicate that the mechanisms responsible for the isotype switch from μ to γ chains and for light chain gene expression are operable at a stage in B-cell development before the expression of surface immunoglobulin and antigen receptors. We suggest that cells containing cytoplasmic γ chains and with the phenotypes described here, reflect early stages in the development of normal B cells, perhaps representing pathways alternative to the major sequence of maturation.

We thank Drs G. Flandrin, D. Buriot, G. Duhamel of Paris, and members of the Southwest Oncology Group, USA, for contributing samples for this study, Dr H. Kubagawa and Mrs S. Labaume for preparing antibodies and Mrs K. Nelson for technical assistance. This work was supported by US PHS grants CA 25408, CA 16673, CA 13148 and DGRST grants 80.7.0220, 78.7.3101. L.B.V. is a Leukemia Society of America Special Fellow and an INSERM Foreign Research Associate.

Received 27 November 1980; accepted 19 January 1981.

1. Vogler, L. B. *et al.* *New Engl. J. Med.* **298**, 872-878 (1978).
2. Cooper, M. D. & Lawton, A. R. in *Cells of Immunoglobulin Synthesis* (eds Pernis, B. & Vogel, H. J.) 411-425 (Academic, New York, 1979).
3. Gathings, W. E., Lawton, A. R. & Cooper, M. D. *Eur. J. Immun.* **7**, 804-810 (1977).
4. Preud'homme, J. L. & Labaume, S. in *In Vitro Methods in Cell-Mediated and Tumor Immunity* (eds Bloom, B. R. & David, J. J.) 155-169 (Academic, New York, 1976).
5. Bollum, F. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4119-4122 (1975).
6. Pearl, E. R. *et al.* *J. Immun.* **120**, 1169-1175 (1975).
7. Brouet, J. C. *et al.* *Blood* **54**, 269-273 (1979).
8. Greaves, M. F. *et al.* *Leukemia Res.* **3**, 353-362 (1979).
9. Greaves, M. F. *et al.* *Leukemia Res.* **3**, 181-191 (1979).
10. Levitt, D. & Cooper, M. D. *Cell* **19**, 617-625 (1980).
11. Siden, E. J., Baltimore, D., Clark, D. & Rosenberg, N. E. *Cell* **16**, 389-396 (1979).
12. Burrows, P., Lejeune, M. & Kearney, J. F. *Nature* **280**, 838-840 (1979).

Natural occurrence but lack of melanotrophic activity of γ-MSH in fish

C. McLean & P. J. Lowry

Pituitary Hormone Laboratory, Department of Chemical Pathology, St Bartholomew's Hospital, London EC1A 7BE, UK

When adrenocorticotropin (ACTH) and β-lipotropin (LPH) are enzymatically cleaved in the pituitary from their common precursor molecule, a glycopeptide with an apparent molecular weight of 16,000 remains (reviewed in ref. 1). Sequencing of a cDNA copy of the bovine pituitary mRNA for the precursor molecule demonstrated that, within the region encoding the 16,000-molecular weight (MW) glycopeptide, there was a short sequence homologous with the α-melanocyte-stimulating hormone (α-MSH) contained in ACTH and the β-MSH within β-LPH². It was therefore proposed that the region of homology coded for a γ-MSH which might have biological activity, although synthetic peptides containing this sequence were shown to have poor melanotrophic activity³. In the pars intermedia further rapid enzyme modification converts ACTH to α-MSH and corticotropin-like intermediate lobe peptide (CLIP), and β-LPH to β-MSH and β-endorphin. Although smaller peptides containing the γ-MSH structure have been detected in extracts of the bovine pars intermedia^{4,5}, the major

post-translational products of the 16,000-MW fragment are present mainly as large glycosylated peptides^{1,4}. We report here that a major peptide product of the neurointermediate lobe of the dogfish pituitary which we had previously shown to lack melanotrophic activity has a γ -MSH-like structure. Therefore if γ -MSH is a peptide with a function it is not one shared with the melanotrophic hormones.

In our previous studies, peptides analogous to mammalian α - and β -MSH were characterized from extracts of the neurointermediate lobe of a more phylogenetically ancient species, the dogfish (*Squalus acanthias*)⁶, the final purification of which is shown in Fig. 1. Peak A was identified as dogfish β -MSH

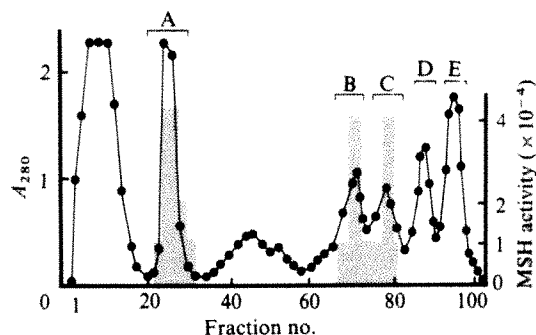


Fig. 1 A_{280} ; MSH activity is shown by the hatched areas. The column (0.6 cm i.d. $\times$ 25 cm length) was developed with a concave gradient running from 20 to 500 mM trimethylamine acetate, pH 5 (total volume 150 ml), and 1-ml fractions were collected every 20 min. A, B and C denote peaks of MSH activity. (From ref. 6.)

whilst B and C were two forms of α -MSH (C was found to be the C-terminally amidated form of B). At concentrations where peaks A and B showed potent MSH activity, D and E were inactive.

Peaks D and E stored from the previous study were repurified by preparative high-pressure liquid chromatography (HPLC, 1×10 cm Spherisorb ODS $10 \mu\text{m}$) developed with a methanol gradient (40–80% in H_2O , containing 1% trifluoroacetic acid). Amino acid analysis of samples of each purified peptide after both acid and enzyme hydrolysis⁷ gave identical compositions: Asn₁, Thr₁, Gly₁, Met₁, Ile₁, Tyr₁, Phe₁, His₁, Lys₂, Trp₁ and Arg₁. E was deduced to be a C-terminally amidated form of D in view of its higher ammonia values and its more basic behaviour during purification on CM-cellulose (Fig. 1). D ($50 \mu\text{g}$) was submitted to automated sequence analysis and gave the unambiguous structure shown in Fig. 2, which also shows the equivalent sequence predicted from the bovine cDNA study and comparisons of the other two species of melanotrophins.

Although there is little sequence homology between D (and its amidated form E) and the γ -MSH bovine sequence, their lack of melanotrophic activity compared with the two other dogfish peptides, α - and β -MSH, and the poor melanotrophic activity of synthetic bovine γ -MSH peptides³ suggest that their assignment as γ -MSH peptides is correct. It is interesting to note here that the unconservative substitutions in the core sequence of dogfish β -MSH compared with bovine β -MSH (Phe for Met and Gly for Glu) have maintained its potent MSH activity whilst both γ -MSHs which have a more recognizable core sequence (Met-X-His-Phe-Arg-Trp) are comparatively inactive as melanotrophic peptides.

The poor periodic acid-Schiff staining of the dogfish (and other fishes, for example, teleosts) intermediate tissue⁸ suggests that the fish γ -MSH/ACTH/LPH precursor is not glycosylated,

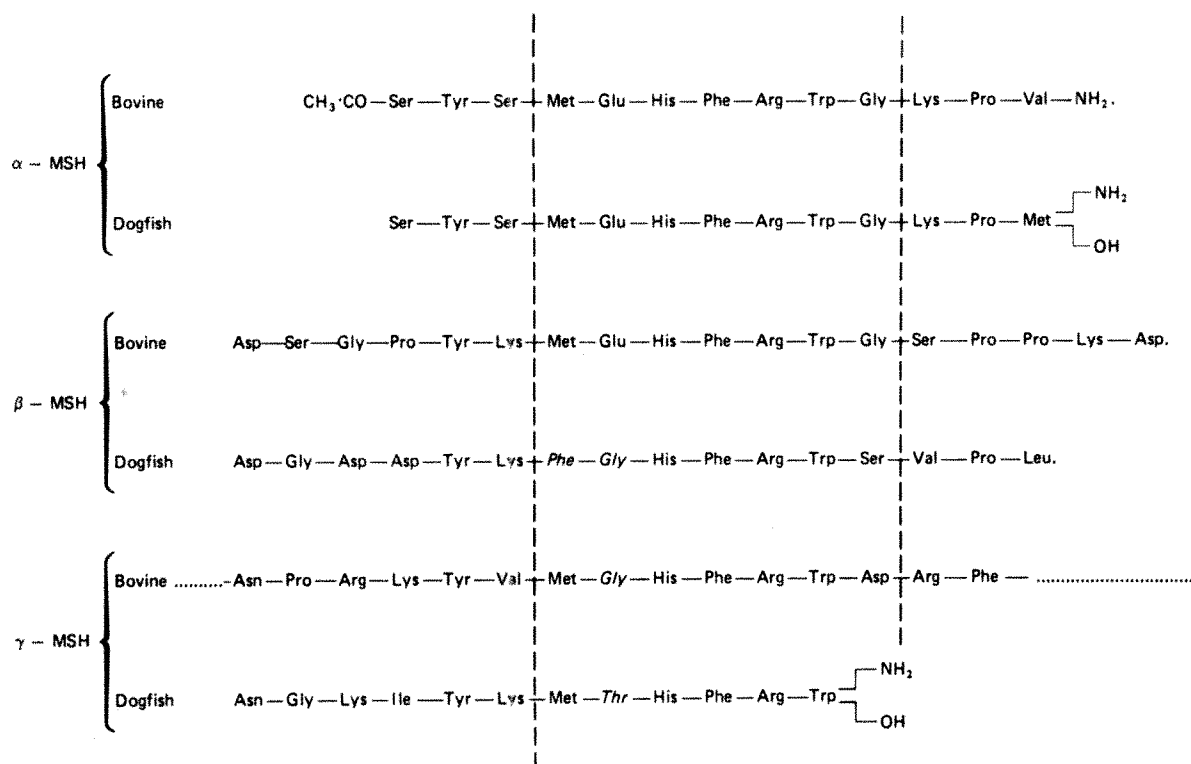


Fig. 2 Comparison of the known and predicted sequences of bovine and dogfish MSH peptides. Sequence analysis of dogfish γ -MSH was carried out in the JEOL-47K sequence analyser using 3 mg Polybrene as carrier. The normal protein programme was used with 0.5 M Quadrol as the reaction buffer. The nitrogen-dried phenyl thiazolinone derivatives were converted to their phenylthiohydantoin derivatives by treatment with $100 \mu\text{l}$ 0.5 M HCl at 70°C for 10 min. After addition of $20 \mu\text{l}$ 1M trisodium citrate the derivatives were directly analysed by HPLC (Altex-ODS, 10–40% acetonitrile/tetrahydrofuran in 10 mM triethylamine phosphate, pH 3, 20 min). Confirmation of samples was carried out by amino analysis after back hydrolysis with HI (150°C , 24 h).

as in the mammals, near the γ -MSH region. It may be this lack of glycosylation that allows the precursor to be processed to the two γ -MSH-like peptides in fish but not in mammals. Their occurrence as major peptide products of the pars intermedia and the C-terminal amidation suggest a biological role for these peptides in the fish but it is unfortunate that the term γ -MSH was coined for such peptides when they have little if any, melanotrophic activity in an animal, like the dogfish, which responds to changes in background by altering skin pigmentation.

We thank the *Biochemical Journal* for permission to reproduce Fig. 1 and Dr J. P. Sumpter for comments.

Received 27 October 1980; accepted 15 January 1981.

1. Eipper, B. A. & Mains, R. E. *Endocr. Rev.* **1**, 1-27 (1980).
2. Nakanishi et al. *Nature* **278**, 423-427 (1979).
3. Ling, N. et al. *Life Sci.* **25**, 1773-1780 (1979).
4. Shibasaki, T. et al. *Nature* **285**, 416-417 (1980).
5. Shibasaki, T. et al. *Biochem. biophys. Res. Commun.* **96**, 1393-1399 (1980).
6. Bennett, H. P. J. et al. *Biochem. J.* **141**, 439-444 (1974).
7. Bennett, H. P. J. et al. *Biochem. J.* **129**, 695-701 (1972).
8. Holmes, R. L. & Ball, J. N. (eds) *Biological Structure and Function. The Pituitary Gland: a Comparative Account* Vol. 4 (Cambridge University Press, 1974).

Slow-reacting substances (leukotrienes) contract human airway and pulmonary vascular smooth muscle *in vitro*

C. J. Hanna*, M. K. Bach†, P. D. Pare* & R. R. Schellenberg*

*University of British Columbia, Pulmonary Research Laboratory, 1081 Burrard Street, Vancouver, British Columbia, Canada V6Z 1Y6

†Department of Hypersensitivity Disease Research, The Upjohn Company, Kalamazoo, Michigan 49001, USA

During a type I allergic reaction histamine, slow-reacting substance of anaphylaxis (SRS-A) and other mediator substances are elaborated from specific tissue sites¹. In allergic asthma these sites are in the lung and the mediator substances cause airway obstruction by contracting smooth muscle and altering mucociliary function²⁻⁴. Unlike histamine, slow-reacting substances (SRSs) have been assessed very little for their roles in obstructive airways disease. This has been partly due to the fact that their chemical nature was unknown until recently and thus pure samples were not available for pharmacological studies. However, SRSs isolated from both immunological and non-immunological reactions have been identified as a combination of two related lipid substances—leukotriene C₄ (LTC) and leukotriene D₄ (LTD)^{5,6}; thus it is now possible to use pure SRSs (leukotrienes) in pharmacological studies of airway smooth muscle. LTC and LTD have been shown to contract guinea pig tracheal and lung parenchymal strips⁵ but there is no evidence that these substances produce similar effects on human lung tissue. To clarify this, *in vitro* pharmacological studies were done to determine the actions of LTC and LTD on smooth muscle strips of human bronchus, pulmonary vein and artery, and lung parenchymal tissue containing smooth muscle components and pleura. As indicated in a preliminary report⁷, all four types of tissues contracted in a dose-dependent fashion to the leukotrienes, although these substances only function as partial agonists.

Lung tissues were obtained from 14 patients undergoing lobectomy or pneumonectomy for isolated peripheral pulmonary lesions. Samples of bronchus, pulmonary vein, pulmonary artery and parenchyma were dissected free and placed in oxygenated Krebs' solution. The bronchial, pulmonary venous and pulmonary arterial samples were then cut into spiral strips 1×4×40 mm; parenchymal strips, with and without pleura, were cut in strips about 5×5×40 mm. These were placed in an organ bath system containing Krebs' solution and drug responses were examined. SRS peaks I and II were obtained from rat peritoneal monocytes by methods described elsewhere^{6,8}. Peak I was identified as leukotriene D₄ (ref. 9) and peak II as leukotriene C₄ (formerly designated LTC-1)¹⁰.

A cumulative drug application procedure¹¹ was used to study the contractile effects of all the agonists, as well as the antagonistic action of FPL 55712 against the leukotrienes, on the four types of lung tissue. All drug concentrations are expressed in molar units. LTD and LTC molarities were determined by converting Upjohn SRS units of activity into moles of leukotriene (LTD: 1 U = 20 pg, molecular weight (MW) = 495; LTC: 1 U × 24 ± 2 pg, MW = 625) and dividing by the dilution volume in litres.

The log dose-response curves for carbachol, histamine (both Sigma), LTC and LTD on human bronchial smooth muscle are shown in Fig. 1. Carbachol and histamine elicited fast onset, sustained, dose-dependent contractions, although the bronchial tissue was consistently more sensitive to carbachol over the full range of the dose-response curve. Carbachol also produced the greatest magnitude of contraction compared to the other agonists. Both carbachol and histamine were active over a wide dose range (5×10^{-9} – 5.5×10^{-4} M) with pD₂ values (calculated as $-\log \text{ED}_{50}$) of 6.301 and 5.658, respectively. Submaximal effects of both agonists were easily reversed when they were washed from the organ baths.

However, responses to the leukotrienes were markedly different: both LTC and LTD caused human bronchial smooth muscle to contract but the responses had a delayed onset and slow development time. The contractions were well sustained (that is, they did not fade) and were dose-dependent. As can be seen in Fig. 1, the bronchial tissue was most sensitive to LTC although this leukotriene produced maximum responses that were only 60–70% that of the maximum responses to carbachol.

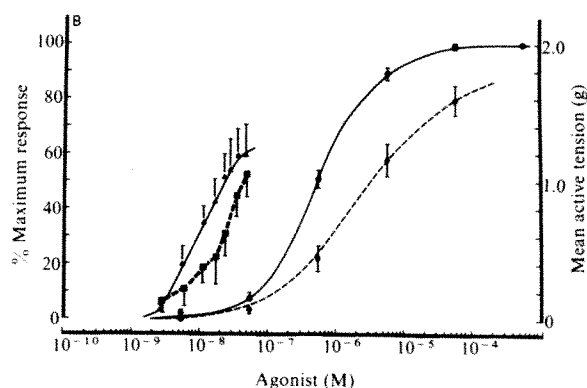


Fig. 1 Log dose-response curves for carbachol (●), histamine (◆), LTC (▲) and LTD (■) on human bronchial smooth muscle. Tissues obtained from patients undergoing thoracic surgery for carcinoma were cut into spiral strips and maintained in 20 ml jacketed organ baths containing oxygenated Krebs' solution, pH 7.4, at 37 °C. Contractions were measured semi-isometrically with Grass FT .03 C force displacement transducers and recorded on a Beckman R511A polygraph. Initial tensions ranged from 1 to 2 g; raw data were collected as active tension (g). Data points are expressed as a percentage of the maximum carbachol response (left ordinate) and a mean active tension scale (right ordinate) is shown for comparison. Each point represents the mean ± s.e.m. of three to seven experiments.

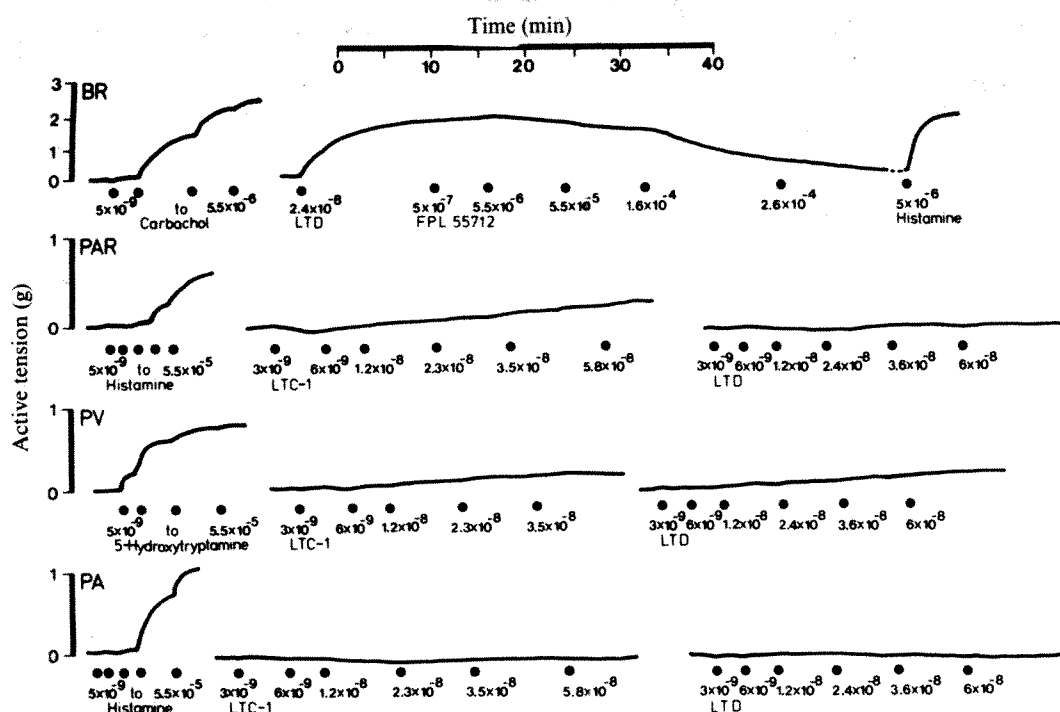


Fig. 2 Original polygraph tracings of the responses of human bronchus (BR), parenchyma (PAR), pulmonary vein (PV) and pulmonary artery (PA) to selected agonists and the antagonist, FPL 55712. Contractile responses are shown in terms of active tension (g) generated per unit of time (min). Drugs were added to the bath in a cumulative manner. LTD, leukotriene D₄; LTC, leukotriene C₄.

LTD similarly acted as a partial agonist. Both the leukotrienes were active over a narrow dose range (3×10^{-9} – 5×10^{-8} M) with pD_2 values of 8.022 for LTC and 7.620 for LTD. Human bronchial tissue was more variable in its responsiveness to LTC and LTD than to carbachol or histamine. It was also found that leukotriene-induced contractions were reversed when these substances were washed from the baths. Furthermore, exposure of tissues to leukotrienes (followed by washing) before addition of another agonist did not alter the tissue responses to the second agonist.

The SRS-A antagonist, FPL 55712 (Fisons), inhibited LTC- and LTD-induced contractions of human bronchial smooth muscle. The upper trace in Fig. 2 shows the response of a human bronchial strip to LTD ($600 \text{ U} = 2.4 \times 10^{-8} \text{ M}$) followed by the addition of FPL 55712. The rate of inhibition was slow and the concentrations of FPL 55712 required to abolish completely the LTD contraction were high ($2.6 \times 10^{-4} \text{ M}$), although the antagonism was clearly dose-dependent. No specific anti-histamine effect was observed with high doses of FPL 55712; the antagonist alone had no effect on the tissues.

The contractile effects of the leukotrienes and other selected agonists on human parenchymal tissue, pulmonary vein and pulmonary artery were also examined to a limited extent. Examples of traces are shown in Fig. 2. In general, the vasoactive amines, histamine or 5-hydroxytryptamine, were more active than LTC or LTD on each of these three types of lung tissue. The parenchymal tissue was slightly more sensitive to LTC (threshold 150 U) than LTD (300 U) and gave a slightly greater maximal contraction (LTC 60%, LTD 35% of the histamine maximum). Pulmonary vein and artery were not as responsive to the leukotrienes as the airway tissues and the responses were characteristically slow in onset and development compared with the amines.

These results suggest that LTC and LTD may act as airway spasmogens when released by an allergic reaction in the human lung. Our data show that human bronchial smooth muscle was very sensitive to the leukotrienes although both substances acted as partial agonists compared with carbachol. Previous studies^{12–14} have shown that human bronchial smooth muscle contracted in response to crude preparations of SRSs; however, our results show that the distinctive 'slow reaction' is caused by the leukotrienes. Dahlen *et al.*¹⁵ have described similar findings with LTC and LTD.

The inability of the leukotrienes to elicit a maximum contractile response, the delayed onset of contraction and the slow rate of FPL 55712 antagonism suggest a mechanism of action on human lung smooth muscle which differs from that seen for other agonists. Such differences indicate that the mechanism of action of the leukotrienes is not closely linked to the contractile apparatus.

In these studies, neither LTC nor LTD had preferential activity on central (bronchial) versus peripheral (parenchymal) airway smooth muscle. As the parenchymal tissues obtained were from patients with usually significant emphysema, the reactivity varied considerably from one tissue to another and we may thus have limited our ability to see a preferential peripheral response. However, this seems unlikely because unreactive tissues failed to respond to all agonists, whereas reactive parenchymal strips not only contracted to agonists but in the same approximate dose range as for the bronchial strips.

These results suggest the possibility that leukotrienes C and D may play a major part in human bronchial asthma.

Histamine and carbachol were from Sigma and FPL 55712 was from Fisons. We acknowledge the support of the Canadian and British Columbia Lung Associations and B.C. Health Care Research Foundation. We thank Drs J. C. Hogg and J. M. Nelems for their contributions and Mrs L. Hare for typing the manuscript. C.J.H. is a Canadian Lung Association Fellow.

Received 20 October 1980; accepted 16 January 1981.

- Piper, P. J. *Pharmac. Ther.* **B3**, 75–98 (1977).
- Kaliner, M., Orange, R. P. & Austen, K. F. *J. exp. Med.* **136**, 556–567 (1972).
- Mezey, R. J., Cohn, M. A., Fernandez, R. J., Januszkiewicz, A. J. & Wanner, A. *Am. Rev. resp. Dis.* **118**, 677–684 (1978).
- Hogg, J. C., Pare, P. D., Boucher, R. C. & Michoud, M.-C. *Can. med. Assoc. J.* **121**, 409–414 (1979).
- Lewis, R. A. *et al. Proc. natl. Acad. Sci. U.S.A.* **77**, 3710–3714 (1980).
- Bach, M. K., Brashler, J. R., Johnson, H. G. & McNee, M. L. *Proc. Symp. SRS-A and Leukotrienes* (Research Studies Press, London, in the press).
- Hanna, C. J., Schellenberg, R. R., Pare, P. D. & Bach, M. K. *Clin. Res.* **28**, 702 (1980).
- Bach, M. K. & Brashler, J. R. *J. Immun.* **120**, 998–1005 (1978).
- Bach, M. K., Brashler, J. R., Hammarstrom, S. & Samuelsson, B. *Biochem. biophys. Res. Commun.* **93**, 1121–1126 (1980).
- Bach, M. K., Brashler, J. R., Hammarstrom, S. & Samuelsson, B. *J. Immun.* **125**, 115–117 (1980).
- Van Rossum, J. M. *Archs int. Pharmacodyn.* **143**, 299–330 (1963).
- Collier, H. O. J. & Sweatman, W. J. F. *Nature* **219**, 864–865 (1968).
- Brocklehurst, W. E. *Adv. Drug. Res.* **5**, 109–113 (1970).
- Mathe, A. A. & Strandberg, K. *Acta physiol. scand.* **82**, 460–465 (1971).
- Dahlen, S.-E., Hedqvist, P., Hammarstrom, S. & Samuelsson, B. *Nature* **288**, 484 (1980).

BOOK REVIEWS

Genes and culture: an awkward hybrid

Richard Dawkins

THIS little book is an attempt by authors who are sophisticated and accomplished in their own field of evolutionary biology to fit human culture into their scheme of things. I therefore opened it with some anticipation. The first chapter is a sensible and persuasive exposition of the claim, which will be uncontested by all who understand what a modest claim it truly is, that living organisms may usefully be treated as machines programmed to propagate copies of the genes responsible for the programming. The authors note the danger that such language

can lead the unwary imagination to a forest glade where tiny pixies work their will by casting spells on inanimate objects, but such good literary fun should not bother the reader who keeps firmly in mind the true aimless nature of DNA and who remembers that the apparently gene-serving behaviour of organisms is an equally aimless consequence of natural selection. Counting on our readers to distinguish the laws of natural selection from the spells of pixies, we will use the image of a purposeful gene to pursue the problem of the origin of learning [p.2].

Ah what touching trust; I cannot resist a wistful sigh. Pixies, hobgoblins, demons and ogres, there will be readers who will find them all lurking in the most unassuming and innocent prose, provided only that "gene" is juxtaposed with language redolent of the boggy world of computers and automation.

Pulliam and Dunford make use of the language of automation in their second chapter, "The Cybernetic Lizard vs. the Toxic Ants". This chapter is part of their long drawn out construction of a niche for learning in the genes-eye view of life. The ethological reader familiar with Konrad Lorenz's classic *Evolution and Modification of Behavior* (Methuen, 1966), or with his later work *Behind the Mirror* (Methuen, 1977), may wonder if the conclusion Pulliam and Dunford reach, important and correct as it is, really needs quite such sustained and protracted rediscovering. Once you start thinking functionally, it certainly makes sense that "the genes, by controlling neural development of the brain, set a learning program that determines how flexible or stereotyped an individual's behavior will be" (p.22). And, moving to the next chapter, "Emotion and Decision", ethologists will need little persuading that evolution of emotions has a lot to do with "genetic selection for decision criteria to

Programmed to Learn: An Essay on the Evolution of Culture. By H. Ronald Pulliam and Christopher Dunford. Pp.138. (Columbia University Press: 1980.) \$10, £5.77.

guide learning. Genes exert control over the behaviour of their survival machines by controlling the development of the neural mechanisms that determine which experiences evoke which emotions" (pp. 26-27). Konrad Lorenz and other ethologists should be given some credit for long ago hammering home the idea that "the genes had to define good and bad for their agent", and that "The solution has been to create emotional sensations within the animal agent which makes it want to avoid 'bad' experiences and repeat 'good' ones" (p. 41).

The authors are now ready to move on to their main subject of culture, which they interpret in the common biological sense of traditional behaviour inherited by non-genetic means. Their chapter "Shortcuts to Learning" makes the reasonable point that, among the "values" which a successful gene might be expected to build into its agents' learning programs, is a predilection for copying the ways of respected individuals. "We might sometimes think to ourselves, 'If they say it is good, it is probably good'." An animal can shortcut the normal learning process by picking up useful tricks from other animals. The Japanese macaque stories give their usual good service in this connection and, mercifully, we are spared the now almost obligatory emphasis on the fact that the great Imo was both female and young when she made her brilliant innovations.

After an interlude on economic exchange theory, we move on to "The Mathematics of Culture", a useful summary of the Cavalli-Sforza and Feldman models of cultural inheritance. There are, of course, immense difficulties in the path of any attempt to apply population genetics-like mathematics to cultural inheritance, but the very difficulties themselves are likely to be illuminating and should not be taken as an excuse for funkling the attempt. The possibility that cultural inheritance might, in some sense, be treated as particulate rather than blending is one that intrigues me. This is my way of wondering if "cultural logic conforms to genetic logic" (Chapter 7). In passing I could not help

noting that Pulliam and Dunford faithfully replicate the notorious cultural mutation in the citation of W. D. Hamilton's "The Genetical Evolution of Social Behaviour" (J. Seger and P. Harvey *New Scientist*, 87, 50-51; 1980). This would have been an especially illuminating example for Pulliam and Dunford themselves to have used since it seems not to be a pure, undirected mutation at all. As Seger and Harvey point out, the "mutant" reading, "The Genetical Theory of Social Behaviour", has arisen three times independently, and is almost certainly a cultural legacy from that most admired of tribal ancestors, R. A. Fisher, and his "The Genetical Theory of Natural Selection".

The authors may or may not establish their anthropological credentials by going to town on "kinship". An impressive few pages bristle with words such as matrilineal and patrilineal. They coin the interesting phrase, "cultural coefficient of relationship", but actually it turns out to be less interesting than I expected it to be. Instead of being the cultural equivalent of the genetic coefficient of relationship, they apply the term to an awkward hybrid between cultural and genetic notions. They seem unable quite to emancipate themselves from the assumption that genetic "fitness" (capacity to propagate copies of one's genes) has some kind of privileged priority over its cultural equivalent (capacity to propagate one's ideas, etc.). Thus, instead of stressing that, in the domain of cultural inheritance, "brothers" may be defined as those (genetically perhaps unrelated) who have sat at the feet of the same master, they come up with what is really only a kind of culturally weighted version of the genetic coefficient.

If a child's parents have different ideas on a particular issue, let us say that a child learns the father's idea with probability P and learns the mother's idea with probability M. Then, the probability that two siblings learn the same idea from the same parent is $P^2 + M^2$. . . This is the probability of cultural identity by descent and can be called the *cultural coefficient of relationship*. Like the genetic coefficient of relationship, the cultural coefficient can be calculated for any pair of relatives.

But if such a cultural coefficient can be calculated for any pair of genetic relatives, can't it also be calculated for any pair of individuals whether genetically related or not?

The last chapter, "The Legend of

Eslok", is rather queer. If I understand it aright, the intention is sensible enough: to use a parable about an imaginary tribe, the Koltinnopinnu to make a point about the conservative role of traditional ritual in preserving ancient wisdom vital for group survival. But the imaginary legend is dragged out with almost Tolkienian attention to gratuitous detail, including learned footnotes on fine points of the fictitious language of the imaginary Koltinnopinno (plurals are formed by

adding "o"). I was left with the feeling that I must be missing the point of the story, since I couldn't see the symbolic (or anagrammatic?) significance of these details. Well, if it comes to that, maybe I missed the point of the whole book: there has to be *some* explanation for my disappointment with it. □

Richard Dawkins is University Lecturer in Animal Behaviour at Oxford University, and a Fellow of New College.

Surprising structures of nucleic acids

C.W. Hilbers

Nucleic Acid Geometry and Dynamics. Edited by R.H. Sarma. Pp.424. (Pergamon: 1980.) Hbk £25, \$55; pbk £11, \$24.50.

IN molecular biology nucleic acids are probably the most discussed of all biological macromolecules. Yet, until the early 1970s, information about the structure these molecules can adopt was mainly based upon results from X-ray diffraction experiments on DNA or RNA fibres. During the past decade the situation has rapidly changed. Several events have contributed to this development. Investigators have been able to crystallize the relatively short transfer RNA and determine its structure to atomic resolution. In addition, newly developed organic synthetic methods have permitted the preparation of well-defined fragments of RNA and DNA, so that single-crystal analysis of these oligonucleotides can be performed, and techniques (NMR, T-jump) for the high-resolution study of the structure and dynamics of nucleic acids in solution have become available.

All of these developments have led to surprising results, outstanding among which is the finding that nucleic acids possess a conformational flexibility that can give rise to structures far more complex than the simple double helix.

Some of the excitement pervading the field is reflected in this book. It is a collection of 16 review chapters by different authors in which the structure and dynamics of nucleic acids are discussed in great detail. The first chapter introduces the nomenclature of nucleic acids, and is followed by a discussion of the way in which structural information can be deduced from NMR measurements and a description of the basics of single-crystal X-ray diffraction techniques as applied to nucleic acids. It is somewhat confusing that although different systems of nomenclature are in current use for nucleic acids and are employed throughout the book, only one is presented in the first chapter. Also, one would have preferred the scaling of the NMR spectra to be given in ppm instead of in Hz.

In the following chapters the basic theme of the book unfolds. Sarma describes the spatial configurations of different forms of DNA, and Seeman gives an account of the crystal structures obtained for oligonucleotides. Lerman's proposal in 1961 that aminoacridine dyes bind to DNA by intercalation between stacked bases has now been verified at atomic resolution for a number of planar dyes intercalated in dinucleotides. Accounts of progress in this area are presented by Sobell and by Berman and Neidle. Rich and his co-workers relate the conformational changes in the sugar phosphate backbone of RNA and DNA fragments, introduced by the intercalation of planar dyes, to the changes occurring in the backbone of tRNA resulting from the intercalation of a base between two consecutive bases.

The intercalation of dyes is taken further by Patel, by Krugh *et al.* and by Crothers *et al.* who discuss the information obtained on this subject from solution studies. Fluctuations in nucleic acid structure as monitored by NMR, ESR and tritium and deuterium exchange techniques are covered by Kallenbach and co-workers and by Lerman and co-workers. Finally, theoretical aspects of the structure, flexibility and accessible surface areas are dealt with in the three chapters by Broyde and Hingerty, Olson, and Alden and Kim.

The book is recommended to anyone involved in structural studies of nucleic acids. It contains a wealth of information, written by experts and illustrated with beautifully coloured molecular models many of which were drawn by Irving Geiss. It is not complete, however, in that it does not contain a chapter devoted to the structure and dynamics of tRNA. A potential buyer should also be aware that a large part of the material has already appeared in *Stereodynamics of Molecular Systems*, edited by R.H. Sarma and published by Pergamon. □

C.W. Hilbers is Professor of Physical Chemistry at the University of Nijmegen, The Netherlands.

Stuck on bacteria

N. Sharon

Bacterial Adherence. Receptors and Recognition, Series B, Vol.6. Edited by E.H. Beachey. Pp.466. (Chapman and Hall/Methuen Inc.: 1980.) £27.50, \$69.95.

IT is a truism that in order to inhabit or infect a susceptible host, bacteria must be able to "stick" to the tissue surfaces of the host. Surprisingly enough, until the early 1970s little attention was paid to the problem of bacterial adherence to cell surfaces, its molecular mechanism and possible role in host-parasite interactions. The recent explosion of activity in this area may be ascribed, at least in part, to the current intense interest in receptors on cell surfaces, and in their role in intercellular recognition. No wonder that the first treatise on bacterial adherence appears as part of the series on "Receptors and Recognition".

The 14 chapters in the book are all written by experts active in the field, including several of the pioneers who for some years worked almost alone on bacterial adherence. Notable among them are J.P. Duguid from the University of Dundee and R.J. Gibbons from Boston. The first contribution, by I. Ofek and E.H. Beachey, is an excellent overview of the field. In the subsequent 11 chapters, emphasis is either on adhesion to different surfaces (mucosal, dental or of plants, for example), on the adhesive properties of specific organisms (Enterobacteriaceae, *Vibrio*, *Neisseria*, for example), or the binding properties of bacterial appendages (fimbriae or pili) and surface polymers (lipoteichoic acid). This loose organization is probably a reflection of the dearth of knowledge of the basic mechanisms involved in adherence. One of the two concluding chapters deals with recognition systems between eukaryotic cells, a topic of relevance to the prokaryote-eukaryote system of bacterial adherence, and the other with prospects for prevention of the association of harmful bacteria with most mucosal surfaces.

From reading the book it is evident that the adherence of many bacteria to epithelial cells is mediated by complementary molecules — adhesins on the bacteria which react with receptors on the cells in a lock and key manner. In the case of *Escherichia coli* and other Enterobacteriaceae, the adhesins are lectin-like molecules in the form of fimbriae that can bind reversibly to sugar residues (predominantly D-mannose) on the surface of the epithelial cells, whereas certain mycoplasmas appear to bind in a similar manner to sialic acid residues. Streptococci, on the other hand, appear to bind to cells via their surface lipoteichoic acid. Much remains, however, to be learned about the nature of the adhesins on other bacteria and about the receptors on the tissues, none of which

has been characterized and about which little is known. More importantly, there is a need for conclusive evidence on the role of adherence in natural infection; if the adhesins play a central role in the ability of pathogenic bacteria to launch infection, then the prevention of adhesion should be an effective way to protect against bacterial infections. This could hopefully be achieved by the application of specific receptor analogues or antibodies to the adhesins.

On the whole, the various chapters are well written with not too much overlap. The references are up to date and the tables abundant and clear. There are many illustrations and, with a few exceptions, they are of high quality, including several artistic scanning electron micrographs

which are most impressive. However, the book suffers from the modern malaise of poor indexing — the number of entries is insufficient and even major items such as adhesins, agglutination, mannose, mucus, virulence and yeast cells have not been included.

This is a valuable book, which should be read not only by researchers in the growing field of bacterial adherence, but also by all those interested in infectious diseases and microbial ecology in general, and all students of cell surface receptors and recognition. □

N. Sharon is the Sadie and Joseph Danciger Professor of Molecular Biology and Head of the Department of Biophysics at The Weizmann Institute of Science, Rehovoth, Israel.

Perspective on the flowering of Fermiology

Walter A. Harrison

Electrons at the Fermi Surface. Edited by M. Springford. Pp.541. (Cambridge University Press: 1980.) £35, \$85.

THIS *Festschrift* is a fitting tribute to Professor David Shoenberg, whose experimental studies of electrons at the Fermi surface span nearly a half century. In 1937, he pioneered the use of the de Haas-van Alphen effect, which arises from the quantized orbits of electrons moving in a magnetic field, to determine the electronic structure of the semimetal bismuth. When such studies were extended to metals, however, they correlated only poorly with knowledge of other properties of the metal and the contemporary theory of electronic structure. The experimental effect tended to "see" the tiniest pieces of Fermi surface, the abstract surface in momentum space (or wavenumber space) which is the

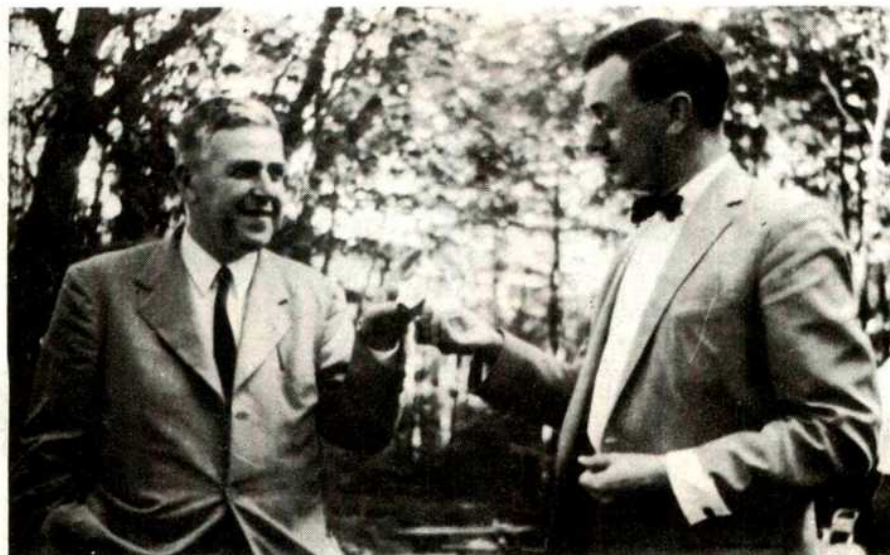
boundary separating electronic states which are occupied in a solid from those which are empty. Most properties, and available theory, focused upon grosser features, so there was little confrontation of theory and experiment. In the early 1950s it was gradually recognized that the de Haas-van Alphen effect was giving unambiguous geometrical information about the Fermi surface, in contrast to properties such as conductivity which depend upon a muddle of geometry, dynamical properties and scattering. The relevant question became clear: what is the *true* Fermi surface? This, rather than: which model of the electronic structure gives the best fit to experiment?

Lifshitz and Kaganov give historical perspective to this period in the first contribution to the *Festschrift*. Their view was of an arbitrary dispersion law for electrons

which was to be deduced from experiment. Indeed, in 1954 Lifshitz and Pogorelov set up a mathematical scheme to deduce the Fermi surface unambiguously from de Haas-van Alphen data, assuming only that the surface was convex and had a centre of symmetry. In 1956, Gunnarsen, a student of Shoenberg, carried this out in detail for aluminium. He needed to sketch in data for directions of field where the effect disappeared but he could then deduce the presence of cushion-shaped segments of Fermi surface. In hindsight this was misleading, since the real Fermi surface consisted of a set of cylindrical tubes; the missing data were different from what he assumed. A study of lead by Gold, another student of Shoenberg, quickly followed and appeared in 1958. Gold came up with similar data, which he interpreted as somewhat larger cushion-shaped segments — these also turned out ultimately to be a set of tubes. However, Gold had other data and he made a nearly-free-electron construction which led to a very intricate, but qualitatively correct Fermi surface. Though he did not have all of the surfaces matched to the correct data, his study pointed the way to the future: predict a Fermi surface from approximate theory and tune the result by experiment. Lifshitz and Kaganov indicate how difficult it was to accept the nearly-free-electron theory until forced to by its many quantitative successes. One can understand why they were not persuaded by the theoretical arguments in its favour by reading the following contribution on many-body effects by Wilkins; a series of customary assumptions can be stated, but success is their principal justification.

The nature of the study of electrons in metals changed at just the time of Gold's work. Pippard had deduced a Fermi surface for copper, published in 1957, using the anomalous skin effect. He also needed a starting picture, which in his case was the volume of intersection of three mutually perpendicular cylinders. In 1958, Shoenberg succeeded in observing de Haas-van Alphen oscillations in a whisker of copper. By 1960, at the famous Fermi Surface Conference in Cooperstown, New York, the multiply connected Fermi surface of copper had been confirmed and the essential correctness of the nearly-free-electron nature of simple metals established. In addition, other techniques, such as the magneto-acoustic effect, had been developed and applied to Fermiology, though the de Haas-van Alphen effect, because of the high order of the oscillations, remained the champion for precision. This flowering of techniques is reflected in the Lifshitz-Kaganov chapter and, most appropriately, in the simple but complete description of the Shoenberg effect in Pippard's contribution.

The subsequent history of the field is also well documented in this book. Once experiment had confirmed and refined the approximations used in band theory, and



David Shoenberg (left) and Brian Pippard discuss a model of the aluminium Fermi surface at the Cooperstown conference, August 1960.

Courtesy of B. W. Roberts, General Electric Research and Development Center

computer technology advanced sufficiently, theory could reliably predict the Fermi surfaces even of transition metals, and ultimately most other properties of these complex systems. The status of this theory is clearly described by Mackintosh and Andersen; indeed, we have the "great calculating machine" postulated by Wigner and Seitz in 1955 which has turned out to be much more interesting than they anticipated. Theory and experiment have moved on to ferromagnetic materials (described in the contribution from Lonzarich), to strained materials (Fawcett, Griessen, Joss, Lee and Perz) and dilute alloys (Coleridge). The de Haas-van Alphen effect has been pushed to extraordinary precision (Coleridge and Templeton) and the waveshapes, as well as the periods, exploited (Higgins and Lowndes). Information about electron scattering times has been extracted (Springford) and can be utilized in transport properties (Chambers).

These are authoritative and well-written articles, with a cohesion arising from the focus on Shoenberg's interests. There is, inevitably, some inconsistency of level and emphasis in such a combined effort and, in this case, an interesting set of repetitions:

the multiple presentations of Fermi liquid theory and of the density functional method reflect the central position which these concepts play in our understanding of metals.

Where will this field go from here? One might have thought that the simple-metal studies had served their purpose by 1960 in leading to pseudopotential perturbation theory which enabled the development of elementary theory of virtually the entire range of metallic properties, and yet that was only the beginning of what Higgins and Lowndes call the "decade of Fermiology". Fermi-surface parameters have now found their way into some physics handbooks so that they become an end in themselves. One could argue that transition metals have been brought into the state of grace of the simple metals and that Fermiology has again served its purpose. However, this field, which is so closely identified with David Shoenberg, seems to carry enough momentum to continue on with no sign of flagging. □

Walter Harrison is Professor of Applied Physics at Stanford University and was involved in the development of the theory of metallic Fermi surfaces.

Confusing collisions

I. S. Grant

Deep-Inelastic and Fusion Reactions with Heavy Ions. Lecture Notes in Physics, 117. Edited by W. von Oertzen. Pp.394. (Springer-Verlag: 1980.) DM49, \$29.

THIS volume contains the proceedings of a symposium held in October 1979 at the Hahn-Meitner-Institut in Berlin. The symposium was chosen as the occasion to dedicate the heavy-ion accelerator VICKSI, which had then been running for almost a year, and to celebrate the centenary of the births of Otto Hahn and Lise Meitner. Hahn and Meitner were associated with the discovery of fission, and it was fitting to devote the symposium to recent studies of the macroscopic behaviour of nuclear matter using heavy ion fusion and deep inelastic reactions.

The experimental papers present evidence for a bewildering variety of different processes which may occur in the collision of two heavy ions. The binary reactions described as fusion fission, fast fission, deep inelastic and quasi-elastic scattering are not clear-cut, but merge into one another in different regions in the space of mass and energy transfer. Nevertheless, it is clear that different phenomena are involved which potentially contain a great deal of information about the dynamical behaviour of nuclear matter. This potential is not easily realized, partly because of the experimental difficulty in recording all the information which may be relevant in each collision, and also because of the obscuring effect of various mechanisms which may contribute to the de-excitation of the composite system or the primary products of the binary reaction. As well as evaporation and fission, there are pre-equilibrium processes which lead to the emission of light particles and which are described in terms (hot spots, jets) which are notions rather than observed entities.

Naturally, the theoretical picture is also confused. One section of the proceedings consists of four theoretical papers all taking a different view of heavy ion collisions. The topic is at the interesting stage when the relationships between its different aspects are not well understood. In a fast-growing field, the proceedings of a symposium held at the end of 1979 are bound to seem a bit dated, but this volume is still a useful survey. □

I. S. Grant is a Senior Lecturer in the Nuclear Structure Group at Manchester University.

Corrigendum

● In the review of *The Principles of Human Biochemical Genetics*, 3rd Edn, by Harry Harris (*Nature* 290, 72) the prices should read: hardback \$65.75, Dfl.135; paperback \$26.75, Dfl.55.

Polemics for palynologists

Peter D. Moore

The Northwest European Pollen Flora, Vol. II. Edited by W. Punt and G.C.S. Clarke. Pp.265. (Elsevier Scientific: 1980.) \$63.50, Dfl.130.

THE second volume in this series will be welcomed by botanists working in such areas as taxonomy, Quaternary palaeoecology, and aerobiology. It follows the first volume in the series in that it consists of a collection of palynological accounts of families which have previously been published in the journal *Reviews of Palaeobotany and Palynology* (Vols 23-28). This practice of reprinting the papers is somewhat inefficient and in future the series will be continued in the form of special issues of the journal dealing specifically with the pollen flora. It is to be hoped that these future volumes will be widely publicized for the sake of those who do not subscribe to the journal and that the special issues will be available separately.

The papers in this volume maintain the high standards set in the first, both in terms of the critical and helpful keys and descriptions, and also the excellent photographs (light and scanning EM). From the palaeoecologist's point of view, particular welcome will be given to the accounts of the Saxifragaceae and Plantaginaceae. In the former, Verbeek-Reuvers takes a fairly conservative position when compared with that of Ferguson and

Webb (*Bot. J. Linn. Soc., Lond.* 63 295; 1970). He emphasizes the problems in distinguishing certain species on the basis of pollen, *Saxifraga aizoides* and *S. oppositifolia* for example. His caution should encourage many palaeopalynologists to go back to the type material and, where necessary, to increase the number of collections of each type to assess the range of variation here described.

In the Plantaginaceae, the differentiation of *Plantago major* and *P. media* is maintained, but the scanning electron micrographs serve only to make one wonder about the criteria upon which it is based. In *P. media*, microechinae (or scabrae) are visible under phase contrast light microscopy, covering the entire exine surface, whereas in *P. major* these are unclear and often totally invisible on the surfaces of verrucae. The SEM pictures of Clarke show that both pollen types possess these scabrae, so it is difficult to understand why they are so obscure on *P. major* under the light microscope.

Other families described in the volume include Solanaceae, Boraginaceae, Valerianaceae, Aceraceae, Haloragaceae and Papaveraceae. It is a book to which all serious palynologists should have ready access. □

Peter D. Moore is Senior Lecturer in the Department of Plant Sciences, King's College, University of London.

2 April 1981

Personal privacy is absolute privacy

How strictly should the rules of privacy apply to the huge banks of personal data now accumulating in the computers of public and private organizations? The European Science Foundation has raised in the past few weeks an intriguing twist on the conventional view that privacy means nothing unless it is absolute. While acknowledging that private persons are entitled to protection, the foundation says that researchers should have access to banks of otherwise confidential data in the interests of "freedom of research". Thus, in its "statement concerning the protection of privacy and the use of personal data for research", the foundation says that governments planning to model legislation along the lines suggested by the Council of Europe's convention on the protection of privacy should be sure to permit access by researchers to otherwise secret data banks. The issue thus raised is intricate, and potentially explosive. The statement in which it is embodied is a much diluted version of an earlier proposal which the General Assembly of the foundation in 1979 would not endorse. Its origins lie in the deep conviction of many social scientists that, as things are, valuable information of potentially great value in the understanding of society is denied to them. But it is unlikely that governments will respond as they would wish to the foundation's proposals. Moreover, there are good reasons why they should not do so.

Of the potential benefits of access by researchers to confidential data, there is very little doubt. In most industrialized countries, publicly compiled data banks would be gold mines for social scientists. Access to computerized health records would make possible far more accurate assessments of the pattern of disease, morbidity and sheer malingering than are possible at present. Access to census data would permit more sensitive studies of marriage patterns, the incidence of social deprivation and of poverty, than the official census takers usually embark on. Access to income tax records would make possible more accurate estimates than those now available of the incidence of wealth within communities — and perhaps even the incidence of tax evasion. In many respects, these data would be of superior quality to that which can be gathered by social scientists themselves using the more traditional methods of survey research. And there is at least some reason to suppose that the benefits would be substantial. More would be learned about the behaviour of different groups within communities than is possible with existing techniques. The frustration of the social scientists that these data are not available to them is understandable. But it is a frustration with which they will have to live for many years to come.

In the past several years, governments have only slowly woken up to the need for protecting personal data held in computer files. Nobody should underestimate the difficulties, the chief of which is the ease with which large quantities of data of a personal character can be stored and then made accessible. In the old days, ordinary people could rely on the difficulty of extracting meaningful data from shelves full of paper files as an assurance of confidentiality. But even where governments have recognized that their citizens are entitled to know that their private affairs will not be broadcast to all who seek to know, problems have arisen in deciding what limits should be placed on the transfer of information from one public authority (say the income tax department) to another (say the police). In equity, there are the strongest reasons why even government departments should be constrained within more or less watertight compartments. It will be a long time before most citizens in advanced societies can be

assured of that. Indeed, outside the United States, enlightened by the Freedom of Information Act, few governments admit the principle that citizens have a right to check data about themselves that may be recorded in the public files. The Council of Europe's convention on privacy is still very much an ideal. It will be an uphill struggle to ensure that it is widely adopted.

This is only one of the reasons why the European Science Foundation's proposal is, at this stage, unhelpful. Rightly or wrongly, it can only raise in ordinary people's minds the fear that supposed guarantees of privacy are not to be taken seriously. And although the foundation's statement suggests that scientists given access to confidential files should be required to follow rules of good conduct, and that they should not make identifiable personal data public without the consent of all those concerned, assurances to this effect will be unpersuasive and legally unenforceable. For where legislation of privacy is contemplated, the onus of responsibility for keeping confidentiality rests with identifiable individuals, the custodians of the sensitive data banks. It is unthinkable that this responsibility could be effectively derogated to whichever academic researchers chose to interest themselves in particular data banks — and unlikely that the researchers would be willing to accept it. The European Science Foundation's proposal is therefore ill-timed.

Scenarios on energy

Last week was a milestone in the short history of the International Institute for Applied Systems Analysis in Vienna, the research institute to which seventeen national academies now provide financial support. For the institute at last published a summary report of its gigantic energy study, on which more than 140 people have been engaged, off and on, since 1973. Readers of the document, *Energy in a Finite World* (Ballinger Publishing Company, \$16.50; Harper and Row Ltd, £12.50) will be of two kinds — those who want to know whether they (or their grandchildren) will be shivering from cold in the year 2030 (when the period covered by the bulk of this study ends) and those who wish to know — perhaps because they have to help find the funds to keep it going — whether the institute's reputation will be enhanced by the appearance of its first important document. On both counts, perhaps in the nature of the questions, the answers are disappointing, even frustrating. As energy scenarios go, the institute's study is more cheerful than most others, but its optimism is so hedged about with qualifications that many fond grandparents will nevertheless think of converting their legacies into string vests and other ways of keeping warm. And the study itself, however impressive its credentials — the leader of the team was Professor Wolf Häfele, a heavyweight in all senses of the word — is deficient in that it fails to acknowledge that the problem of energy is no longer technical but political.

It is now conventional to say that there has been an energy problem since the second half of 1973, when the members of the Organization of Petroleum Exporting Countries (OPEC) abolished by the stroke of a pen what had been until then the most valuable of the developed world's energy resources — crude oil at a price of \$1.50 a barrel. (The same oil costing \$32.50 a barrel seems a very different commodity to those who use it.) The institute's study, urging that those preoccupied with the energy

crisis should take a long view, avoids the trap of suggesting that the energy problem came of age in 1973. The truth is, of course, that it has been with us always. Was not the case for nuclear power in the 1950s based on the assumption that oil reserves would be exhausted in a couple of decades or so? Did not the great W.S. Jevons, leader of the Manchester Liberal School in the 1860s, bemoan the condition of Britain when the coal ran out? For the past 4,000 years, since the first machines were invented, communities of different kinds have found themselves in difficulties for lack of the means for driving their machines. If there is anything special about 1973, it is that so many complacent governments of oil-importing states were prepared to believe that oil-producing states would go on indefinitely selling their oil in whatever quantities were demanded and at a knockdown price. This perspective is a plus for the institute.

So what is special about the present period? The institute's report deals with the period 1975–2030 and starts with a daring but realistic assumption — population growth in the world as a whole during that period will be faster than in any previous time, but will afterwards begin to level off. The assumption is daring because most comparable studies have assumed that population growth will continue and be what is loosely (and falsely) called "exponential". Many doomsayers will complain that the institute is hopelessly optimistic. But the assumption is realistic for at least two reasons. Demographic history shows that sooner or later all communities learn how to balance births against deaths (and the doubling of world population in the next 50 years is less than the threefold increase in Britain between 1815 and 1865, when that demographic transition was more or less accomplished). The second reason is that those states that fail to keep their populations within bounds will end up in catastrophe. This is another plus for the institute.

So too is what the institute has to say about a number of other features of the business of energy forecasting. Technology will indeed make economic activity less energy-intensive than at present. Energy consumption per unit of Gross Domestic Product will continue to fall in developed countries as prosperity (Gross Domestic Product per head of population) continues to increase. And developing countries will not repeat the excesses of the states now industrialized, with the result that they will be able to attain (in half a century) prosperity not very different (outside Africa and South-East Asia) from that now taken for granted in Western Europe without ever having to consume more than 1 watt of energy for each annual dollar's worth of product. These assumptions obviously take the edge off the calculation of how supply and demand for energy can ever be matched. Once the need to worry about what will happen when developing countries consume energy as intensively as developed countries do now is done away with, the collective pressure on resources is diminished and the problem of supply and demand becomes manageable. This part of the institute's argument is well documented and well argued. Only those with a vested interest in believing that the worst will always happen will protest. So this is very much a plus for the institute.

How then does the future look? That is the question futurologists inevitably ask. Mercifully, the institute answers the question in the right spirit, thereby avoiding the snares which have trapped its predecessors. The flaw in most studies of this kind is that a scenario, while not a prophecy, is a hint of what may happen "other things being equal". The institute's study is more imaginative. The objective of a scenario, its study might have said explicitly, is not to show what the future will be like — for none of us can know — but to show that, on plausible and realistic assumptions, the future is possible. The problem, in other words, is to prove what the mathematicians would call an existence theorem. When that is done, extraneous influences — market forces for one, ingenuity for another — will do their work. And once, it has been demonstrated that getting from A (here) to B (there) is possible, it is best not to try to predict how it will be done. Apart from the safe prediction that for most of the next half-century fossil fuels (oil included) will continue to predominate, and from the guesses that global energy production will increase

dramatically in the next 20 years (by a factor of between 1.6 and 2.0) and fantastically (by a factor of between 2.6 and 4.2) by 2030, the institute gives few hostages to fortune. This is the good sense to be expected of a bunch of self-confessed natural scientists. Yet another plus for the institute, it will be concluded.

But how do we get from A to B? That is not a question in futurology, but a question for the present. What do we do now to make the possible future happen? The institute's report takes several novel and stimulating tacks. It has a cheerful account of the virtues of hydrogen as a convenient alternative to electricity as a secondary fuel (and urges that there should be more deliberate attempts to trap the hydrogen released from the deformation produced in the roots of plants in associations with nitrogen-fixing bacteria). The report is lukewarm but not frigid on biomass, futuristically open-minded about energy-collecting satellites in synchronous orbits and philosophical about nuclear power in both its forms — fission and fusion; it's a matter of time, and of giving people a chance to believe that those technologies can be made safely to work. Hurrah, it will be said; maybe there are several routes from A to B. That, for the institute, is a plus of a kind — perhaps half a plus. For these imaginative (but well written) parts of the summary report do not tell us what to do to get from A to B. Putting money into biomass (another name for wasteful forestry) or hydrogen research is only half an answer. For the truth about energy has become political, and will remain so at least until 2030.

Internationally, the proof that energy is a political matter came in 1969, with the *de facto* revocation by Libya of the licences it had earlier given to the multinational oil companies operating there; OPEC's resolve (in 1973) to put flesh on the bare bones of its long-declared intention to secure a fair share of the rent (in Adam Smith's sense) for its oil followed not long after. Domestically, until this year, the United States has been the worst offender, for political reasons charging too little for gasoline and other petroleum products, in the process giving its citizens the illusion that oil is rather like atmospheric oxygen — breathe in (exert demand) and the vacuum will be filled. President Reagan in the past two months has succeeded where his three immediate predecessors (Carter, Ford and Nixon) failed in deregulating the price of natural gas and crude oil in the United States. The short-term consequences have been dramatic — United States imports of petroleum have fallen to just over 3 million barrels of oil a day — 60 per cent of what they were two years ago. Even so, the prices of oil products in the United States are still less than elsewhere — no American administration, not even Reagan's, would risk a thoroughly free market (freedom for exporters and what that entails). The Americans, however, are only the worst and not the only offenders. The British, having found oil in the North Sea, are busily using the tax revenues arising as a poultice for persistent economic ills. Meanwhile, the moderation of OPEC in the past few months owes as much to the continuing war between Iran and Iraq, and the strengthening military relationship between Saudi Arabia and the United States, as to an explicit decision about the price oil can now command. Nobody is much concerned about Mr Willy Brandt's new-found constituency, the developing countries that happen to have no oil.

The institute's report acknowledges that such extraneous factors must influence the pattern of energy consumption in the decades ahead. Its good cheer is based on the assumption that people will behave rationally, on energy pricing, on the development of nuclear power, on a flexible strategy for research on alternative sources of energy, and in the avoidance of war. The caveats are sensible. It follows, however, that what the world now needs is not an existence theorem but a prescription for the avoidance of irrationality. By its nature, the Vienna institute is ill-equipped to help. Its members are, after all, predominantly natural scientists. *Energy in a Finite World* will bring credit to the institute — and would have brought more if it had seen the light of day some years ago. With the passage of the past eight years, however, the problem of energy has been transformed from one of technical analysis to a set of problems in politics, international and domestic. Who will provide the handbook for their solution?

Congress writes back Reagan economies

Democrats prepare for battle

Washington

Members of the US Congress are digging in for a period of extended skirmishing over Mr Reagan's proposed cuts in federal support for research. Democrats are already fighting to restore many of the cuts, and in the House of Representatives at least — where they still have a majority — have already had some success. In contrast, Republicans are sticking as tightly as they can to the fiscal surgery recommended by the Office of Management and Budget (OMB).

Last week, the science and research subcommittee of the House Committee on Science and Technology voted to authorize a budget of \$1,161 million for the National Science Foundation (NSF) in the fiscal year 1982, 12 per cent more than Mr Reagan's proposed budget of \$1,033 million. As expected the committee voted to restore funding for science and engineering education. Mr Reagan's budget proposed that NSF involvement in these activities be terminated; but the subcommittee has restored \$75 million, close to the figure proposed by Mr Carter. It also agreed to restore \$26.1 million for research in the biological, behavioural and social sciences, another target of Mr Reagan's pruning.

Not all the subcommittee decisions went in the same direction. The members also approved a recommendation from Mr Allen Ertel (Democrat) to eliminate \$12 million which had been allocated for NSF's new Ocean Margin Drilling Program, initiated last year as a joint venture with the oil industry. A similar proposal to eliminate ocean margin drilling activities made by the House last year was overruled later by the Senate.

The subcommittee's recommendations on the NSF budget now go to the full science and technology committee, which will decide whether or not they should be passed intact to the floor of the House and where further opposition from Republican committee members is expected. The previous enthusiastic support for NSF by the full committee is now less certain — some Democrats may vote with the Republicans (as they have already done to push through proposed reductions in earthquake research and other projects sponsored by the Federal Emergency Management Agency).

In the Republican-dominated Senate it is generally expected that action on the NSF authorization will be held to Mr Reagan's level. If so, the major debate will come when the House and the Senate meet to

reconcile their approaches. These proceedings could be repeated when Congress decides how much money it wants to hand over to NSF rather than just to authorize.

A similar confrontation is brewing over the budget for energy research. President Reagan proposes major cuts in support for research into alternatives to oil and fossil fuels, such as solar and geothermal energy and alcohol fuels, arguing that these will be encouraged by proposed tax credits and that demonstration projects represent a hand-out to the private sector; in contrast he has recommended a substantial increase in funding for nuclear energy.

House Democrats are trying hard to reverse these proposals, sometimes with the assistance of Republicans who accept the economic rationality of federal investment in energy-saving technologies. The energy development subcommittee of the House Science and Technology Committee has already voted to increase the solar research budget from the \$193 million proposed by Mr Reagan to more than \$500 million, closer to Mr Carter's figures.

This proposal, if sustained by the full House of Representatives and if the Senate sticks with Mr Reagan's policy, is again likely to create substantial conflict. Aware of this possibility, the Department of Energy is said to be drawing up new legislation which would dramatically reduce its responsibility for developing new and renewable energy sources.

Despite Republican predictions of economic doom should the budget cuts not be approved by Congress — and threats of a presidential veto in this eventuality — many individual researchers remain confident that congressional allies will help save their projects. One example is research

in magnetohydrodynamics (MHD), aimed at improving the efficiency of coal-fired power stations, on which more than \$400 million of government funds has already been spent.

In its budget proposals the Reagan Administration suggests that research and development on MHD, previously scheduled to receive \$60 million in 1982, be terminated, arguing that the demonstration of engineering feasibility should become the responsibility of the private sector. A senior OMB official explained that the agency had "agonized greatly" over whether to make the cut; but that although MHD research met two of the government's present criteria for support — high-risk research with a long lead-time — it did not meet the third — a high probability that any successes would be taken up by private utilities.

MHD engineers, however, claim that many utilities remain committed to MHD, which is predicted to increase the efficiency of electricity production from coal by 50 per cent. They point out that cuts at this stage would not make sense, as most of the money is required to cover the operation of two plants whose construction is nearing completion, at the University of Tennessee's MHD Coal-Fired Flow Facility in Tullahoma, and the Component Development and Integration Facility for which superconducting magnets are being provided through the Massachusetts Institute of Technology.

Officials in Tennessee said last week that they expected local congressmen such as Representative Albert Gore or Senator Jim Sasser to argue successfully that the funds should be restored. In other years, their confidence might have been justified. This year, with OMB defining demonstration projects in energy research as "the new

Hidden cheer culled from statistics

Washington

In a neat display of the power of statistics, Reagan Administration officials have estimated that, even with the proposed cutbacks, spending on basic research will increase in real terms in the fiscal year 1982. Figures prepared by the Office of Management and Budget, assisted by staff from the Office of Science and Technology Policy, predict that funds for basic research will grow by 1.5 per cent from the 1981 level. At the National Science Foundation, where the cuts have focused on science and education and social and economic research, basic research in the natural sciences would show 3 per cent real growth.

The catch to these otherwise heartening figures is that they are based on an officially-predicted inflation rate of 7.5 per cent — an optimistic forecast that

assumes almost instantaneous success for Mr Reagan's economic policy.

Overall, the Office of Management and Budget expects the research and development budget for 1981 to be "a few hundred million dollars" higher than the previously expected figure of £35,600 million, largely because cancellations in civilian research will be more than offset by a \$500 million increase in military research. The net result of Mr Reagan's proposals would be to increase research spending by about 13 per cent between 1981 and 1982, to a total of \$40,600 million. Compared with Mr Carter's proposals, the research budget for the Department of Defense would be increased by an extra \$1,500 million, while the non-defence budget would be \$2,200 million less than first proposed.

David Dickson

pork barrel", there is more doubt.

Last week, for example, even the chairman of the House science and technology committee, Mr Donal Fuqua (Democrat) failed to persuade the members of the energy supply subcommittee to support his efforts to reinstate funds for a synthetic fuels demonstration project which faces shutdown without further federal support. Other areas in which efforts to restore funds are being made are for the Ocean Thermal Energy Conversion programmes — OTEC — described by OMB officials as being of very limited pay-off — and the decision to hold back on establishing a Center for Fusion Energy and its Fusion Energy Device, both centrepieces of Mr Carter's plans for the rapid development of magnetic confinement fusion.

David Dickson

UK aerospace

Eager for orders

British industry is beginning to wake up to space communications as a commercial proposition. British Aerospace, the largest British company with a space interest, is on tenterhooks, waiting to hear whether it has won two major contracts for communications satellites. The orders at stake are for a British defence communications satellite network and for the Arab regional satellite communications system, Arabsat.

The contracts would be a watershed for the company, for it has hitherto had only the European Space Agency as a customer for satellite systems. If its bids are now successful, it hopes to compete for similar orders in other parts of the world.

Arabsat, which is also being contested by Ford Aerospace and the Hughes Corporation, is the most crucial of the two. British Aerospace is cheered that it has appeared on the short list of three out of the five initial bidders. The Arab countries are to make a decision within four or five weeks.

Competition for the defence satellite system is not so fierce. The British government, which had been expected to award the contract at the end of March, is taking its time choosing between British Aerospace and Marconi. The plan is to order three satellites, worth a total of about £100 million, one of which will be operational at any time and which will replace the seven-year-old Skynet satellite. The company that wins will hope to gain a foothold in the NATO market.

British Aerospace's attempts to become a major supplier of telecommunications systems have also extended to setting up joint ventures with other companies. Together with Plessey, which has built satellite earth stations, it says it can offer a team capable of building complete defence communications satellite systems. It is also planning a joint venture with the US Comsat General corporation to lease military satellite communications services. And its ambition to compete for more

contracts like Arabsat led it last month to set up a group called Satcom International with the French company Matra. Satcom plans to put in a bid for a regional satellite for Australia in May.

British Aerospace is selling itself on its experience. It has played a leading part in the telecommunications activities of the European Space Agency, having been the prime contractor for the series of Orbital Test Satellites and the more sophisticated European Communications Satellites. It also prides itself on being the only non-American company to contribute to Comsat General's Comstar programme for long distance telephony within the United States.

Attempts to exploit the commercial potential of space may have come rather late in Britain. But the present flurry of activity is not confined to aerospace and electronics companies. The Department of Industry says that an increasing variety of organizations, from banks to pension funds, now think of telecommunications when asking where to put their money.

How matters will develop depends not only on the industry's ability to compete and fulfil its promises but also on what the British government decides. The industry is now in limbo waiting for a government response to a confidential report by the Central Policy Review Staff on the commercial exploitation of space and for the publication of a Home Office report on direct broadcast television by satellite. The reports, if nothing else, may recommend making it easier for venture capital and private enterprise to take a share of the action.

Judy Redfearn

UK nuclear power

Battle begun

The first shots have been fired in the impending battle about the building of a pressurized water reactor (PWR) in Britain, even though the Department of Energy has not yet decided when the proposed public inquiry will be opened. Last week, the Friends of the Earth (UK branch) issued a document subtitled "A critique of the government's nuclear power programme" (Friends of the Earth, London; £1.50) which concludes that its

own case against the proposed reactor, to be built at Sizewell in Suffolk, on the North Sea coast, is so strong that a public inquiry is not necessary.

The argument is straightforward. First, forecasts of electricity demand in the United Kingdom, revised downwards for the past several years, are "wholly unrealistic" and are likely to be further undermined by changing patterns in the demand for electricity. Second, the cost of building the first British pressurized water reactor is likely to be twice the Electricity Generating Board's estimate (of £850 per kilowatt) while the cost of generating electricity will be magnified by the alleged low availability of reactors based on the Westinghouse design.

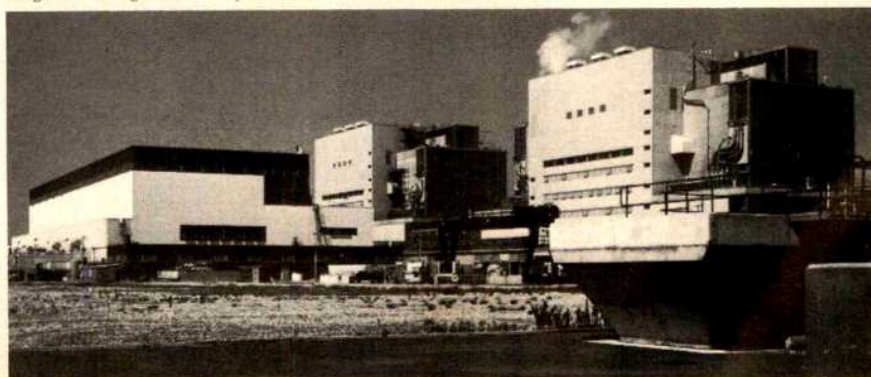
Third, the report (which includes a helpful account of the design changes required by the British nuclear inspectors) says that "the unique hazards of the American PWR must not be introduced into Britain". It is especially critical of a study by the nuclear inspectors of the generic safety of this type of reactor, which has not been published in full. The report also argues that "unresolved" problems concerning the safety of these reactors have not yet been resolved.

The conclusion is that Britain (and British taxpayers) would be better advised to put their money into renewable resources. Insulating lofts to a high standard would be six times as productive as building a pressurized water reactor at Sizewell, according to the Friends of the Earth.

Earlier this week, the Central Electricity Generating Board — the chief but not the only target of the Friends of the Earth — had not decided whether, and if so how, to reply to the criticisms levelled at it. To the extent that some of the complaints coincide with the criticisms of the House of Commons Select Committee on Energy some weeks ago (see *Nature* 19 February p.621), the authors of the new report may have to wait until the formal response from the Department of Energy.

The intended public inquiry about the proposal to build a single pressurized water reactor at Sizewell is now a formal commitment by the British government. Although the form of the inquiry has not yet been settled it is likely to be generic, dealing not with the characteristics of the Sizewell site

Magnox at Dungeness A — open soon?



but with the virtues (or lack thereof) of pressurized water reactors in general. Most probably, the inquiry will not now begin until the second half of 1982, implying that construction of the reactor system cannot begin until late in 1983.

For the electricity undertakings, the chief objective of the proposal to adapt an American design to British circumstances is the hope that pressurized water reactors could be built more quickly than indigenous gas-cooled designs. The Friends of the Earth document, however, is sceptical of the six-year construction time on which the generating board is counting.

Meanwhile, the Central Electricity Generating Board is somewhat encouraged that four of its oldest gas-cooled reactors, shut down some months ago after the discovery of cracks in the expansion joints of the gas-cooling circuits, can soon be started up again. The first of these reactors, at Berkeley on the Severn, is to be restarted in a few months following a demonstration that the cracks have not grown since the reactors were first built.

Czech nuclear power

Action on safety

Czechoslovakia's nuclear energy authorities have recently made overtures to their British counterparts, suggesting that the agreement on cooperation in nuclear safety, signed almost 15 years ago, should be revitalized. In particular, the Czechs are seeking an exchange of information and expertise on the disposal of nuclear waste and on the training of power station personnel.

The human factor in nuclear accidents has lately become a major concern in Czechoslovakia. Last August, after repeated denials, and without any attempts to prepare public opinion, the authorities suddenly admitted that in the past few years there had been two major accidents at the Jaslovské Bohunice nuclear power station.

During a recent phone-in programme on Prague radio, Dr E. Medke, director of the Nuclear Power Station Research Institute at Jaslovské Bohunice, described the institute's "high quality" research programme, aimed at the "long-term, safe, reliable and in particular, economically effective" operation of Czechoslovakia's nuclear power stations. This, he said, is to include work on the human element in safety. Analysis of nuclear accidents throughout the world, he said, shows that only 20 per cent were caused by equipment failure and the rest by operator error.

A training programme is under way to provide the 1,300 graduate engineers whom the country's nuclear programme will need by 1990. After taking a first degree in engineering, the candidate will go on to a two-year theoretical course, and then to a power station for practical

training. This will include work on a simulator, now being built at Trnava. After careful screening, the student will then become an assistant at a nuclear power station, and will finally take examinations set by the State Commission to qualify for a nuclear engineer's licence.

Czechoslovakia, which is more than self-sufficient in uranium, has a major commitment to nuclear energy, and accordingly tried to hush up the accidents at Jaslovské Bohunice, one of which involved two deaths. But in 1978 the Charter-77 movement published a detailed report on the incidents — and although then still unprepared to acknowledge that anything untoward had occurred, the authorities immediately inaugurated a policy of recruiting nuclear power station workers of a better psychological calibre — a major recommendation in the Chartists' report.

In the radio programme, Dr Medke made no mention of the accidents, and left his deputy, Dr S. Novak, to deal with public apprehensions about nuclear power. Regarding the station now under construction at Mochovce near the Hungarian frontier, Dr Novak said that there has never been a case in which one country has contaminated another's territory by nuclear effluent, and that all Czechoslovak power stations have adequate safety measures to prevent environmental damage in the case of breakdown or failure. This statement contradicts the Chartists' report which alleges that in one of the incidents at Jaslovské Bohunice, effluent reached a watercourse. It also contradicts claims from Bulgarian environmentalists that radioactive contaminants can be found in the Danube — although these are conventionally assumed to be of "capitalist" origin.

Vera Rich

Soviet space programme

Mongolia's turn

Last week's spaceflight of the Mongolian cosmonaut, Jugderdemidyn Gurrachaa, coincided with the start of the diamond jubilee celebrations of the Mongolian People's Republic. A few days before the launch, Mongolia's new Five Year Plan was published, extolling the value of science and technology to the development of the republic.

Not surprisingly, most of the experiments performed by Gurrachaa and his Soviet colleagues aboard the Salyut-6 were aimed at practical applications rather than basic research. One of the major tasks was the mapping of Mongolian natural resources, using conventional cameras and a Bulgarian "Spektr-15" multichannel spectrophotometer. This survey will provide data for existing joint Comecon research projects on Mongolian resources including the irrigation of the Gobi Desert. A programme of simultaneous photo-



Gurrachaa — applying science?

graphy of the Salyut spacecraft from Mongolia's two satellite tracking stations (in Ulan Bator and the Gobi) was intended to provide data on factors affecting space geodesy.

The biological programme involved the monitoring of the cosmonauts' biorhythms, work capacity, cardiovascular function and psychological response to the conditions of space travel. The biorhythm observations were started in 1977 during the selection of candidates for cosmonaut training — apparently the first time that a psychological and biochemical study of Mongolian citizens had ever been undertaken on a daily basis. Technological experiments included the study of natural quartz and of various alloys of copper under zero gravity.

Fundamental research was not completely ignored in Gurrachaa's mission. An experiment codenamed in Russian "Izlučenje" (radiation) continued the Mongolian work on primary cosmic rays with energies in the TeV range which was begun in 1972 with an experiment designed by S. Sugar and B. Chadraa of the Laboratory of High Energy Physics of the Academy's Institute of Physics for the joint Comecon Interkosmos-6 satellite in 1972.

Vera Rich

UK computer industry

ICL falters

International Computers Limited (ICL), Britain's largest computer manufacturer, announced last month that it has been trading at a substantial loss since the middle of 1980. The news of ICL's problems came only a few weeks after similar news from Imperial Chemical Industries (ICI). Until the beginning of this year, both companies had been regarded as profitable examples of Britain's otherwise ailing industry.

Unlike ICI, however, ICL is to be bailed out by government money. Sir Keith Joseph, Secretary of State for Industry, has said that the government will act as guarantor for up to £200 million worth of

loans which the company is expected to secure from the banks. Help is also to be provided for ICL's research and development programme which last year consumed about 8 per cent of the company's £715 million turnover.

Sir Keith's intervention is seen as a major blow to the Conservative government's policy of not interfering in private industry. But the decision was taken only after several months of searching for private investors had found no takers. The chief justification for supporting ICL when other private industries are allowed to go to the wall seems to be that the smooth operation of many government offices depends heavily on its computers.

ICL is something of a favourite of British governments. When it was created in 1968 from the amalgamation of two computer companies, the government retained a substantial shareholding and supported it heavily during its early years both with cash and with awards of almost all contracts for large computers for government departments. In the 1970s, the company became profitable and in late 1979, Mrs Thatcher's government sold off its shares. But only 18 months later, ICL's profits took a lurch.

Precisely why the company's fortunes changed so suddenly seems to be a mystery. The company itself says that the world recession combined with the strong value of the pound and high interest rates is responsible. There were hints of trouble, it says, in the first half of 1980 but during the second half both home and overseas orders fell off markedly. Others have said that poor management, a concentration on large computers when the market is moving towards desk-top machines, spending too much on parts of the business not connected with manufacturing and even underpricing goods, are also responsible.

Neither the government nor the banks, however, seem to have stipulated conditions for the £200 million loan. The company hopes to ride out the recession by continuing business much as before and economizing wherever possible. It says, for example, that its research and development programme will continue at much the same level as last year, with the same mix of research on large and small machines. One of the projects on which it pins hopes for the future is the large Distributed Array Processor which has so far been sold to the British National Oil Corporation, for seabed exploration, and Queen Mary College University of London. The Science Research Council is considering buying one for its Rutherford or Daresbury laboratories.

Sir Keith Joseph hopes that the government will not have to make good its guarantee, believing that the loan will see ICL through the recession until business picks up again. Meanwhile, the net is still being cast far and wide in the hope of finding private investment on suitable terms.

Judy Redfearn

London medical schools

Mergers galore

The University of London has made a modest beginning on the reorganization of its twelve medical schools. Last week, the senate of the university approved a plan to save something like £5 million a year, or ten per cent of the present cost to the university of medical education. Some two-thirds of the saving will come from a general reduction of staff-student ratios throughout the medical schools.

The accompanying reorganization of the London medical schools is only a pale shadow of the grand design put forward more than a year ago by a committee under Lord Flowers, rector of Imperial College. At the senate meeting last week, it was formally agreed that the Westminster Hospital Medical School (a clinical school now recruiting about 100 students a year) should be merged with the medical school at Charing Cross Hospital (now located in West London). Other proposed groupings of medical schools have been given until later in the year to work out how they will coordinate their teaching.

The Westminster school considers itself unfairly singled out by these developments. Ever since the publication of the Flowers Report, Westminster has been urging the benefits of small medical schools, complaining that its students or their successors will be lost in a medical school with an annual intake of 180, the planning figure for clinical students at Charing Cross. It also says that the efforts that have been made in the past year or so to coordinate teaching between the two hospitals may be jeopardized by the shotgun marriage now decreed. Given forward commitments to students, it may be as much as five years before teaching comes to an end on the Westminster site.

In retrospect, the schools least affected by the changes now proposed are those at Guy's Hospital and St Thomas's Hospital, which announced last year that they would form a joint school, with a single governing body, but would continue to operate on both the present sites. This arrangement has now been sanctioned by the university senate. The other consortia from which the university is seeking financial savings are University College, Middlesex and St Mary's, which are required between them to save the equivalent of one clinical and one preclinical school; St Bartholomew's and the London Hospital, which have been given until June to suggest how they will merge; and King's College (which has a preclinical school) and King's College Hospital, which seem to be willing to fall in with the recommendation that they should form an integrated unit.

The proposals accepted last week by the senate were recommended by the Joint Planning Committee on Medical Education, set up to pilot the Flowers Report through the university. In the past

year, and especially as a result of a detailed analysis of the costs of the London medical schools carried out by a firm of accountants, it became clear that changes of staff-student ratios offered the most rapid route to cost reductions. The committee accepts as a standard for future staffing a ratio of 1:10 for preclinical schools and 1:7 for clinical schools. This would mean the loss of 180 academic posts, which it is hoped can be achieved by natural retirement and wastage in the next five or six years.

Running through the report of the joint committee is the view that medical education should be concentrated in larger units than at present, especially where modern buildings have been provided. In this spirit, the committee recommends that the development of the St George's Hospital site in South London should be continued. The future of the postgraduate medical institutes remains entirely open, however. With the exception of the Institute of Dermatology (which is to merge with St Thomas's Hospital Medical School), the institutes have been seriously affected by the reductions in the past year in the numbers of overseas students. The university is hoping that central government will help bail them out, but no decision has yet been reached.

Last week's decision by the university senate is likely to be regarded, within the University of London, as some kind of precedent for the reorganization of the non-medical parts of the university likely to be recommended by the Swinnerton-Dyer committee, due to report in September, and elsewhere as a sign that, if pressed hard, even British universities can change their ways.

Law of the Sea

US backtracks

Washington

The Reagan Administration's decision to carry out a complete review of the US position on the draft Law of the Sea Treaty, now under discussion at the United Nations in New York, indicates that significant new attitudes may emerge towards international resource disputes ranging from the surface of the Moon to the minerals of the Antarctic continent.

The United States is considering whether to adhere to the concept that natural resources which lie outside the immediate political domain of individual nations should be treated as "the common heritage of mankind" — a term coined by the United Nations — or whether, in the name of free enterprise and the national security implications of a shortage of strategic minerals such as cobalt and manganese, this concept should be abandoned, together with attempts to set up international regulationsto control exploitation.

A new twist to the Law of the Sea negotiations comes from the likely take-over of the Kennecott Corporation, a leading critic

of the United States' negotiating position, by Standard Oil of Ohio (Sohio). Sohio is 53 per cent owned by British Petroleum, which in turn is 45 per cent owned by the British government. In the past, BP Minerals, a member with Kennecott of one of the five consortia interested in deep-sea mining, has been keen to finalize the treaty to protect its existing exploration and mining operations.

In theory, a Sohio/Kennecott deal could favour a successful outcome to the Law of the Sea negotiations. In practice, the outcome could depend more on whether the Reagan Administration listens to arguments from international companies that a treaty is essential to the security of their future trading operations, or whether it gives greater weight to conservative pressure groups arguing that the whole "common heritage" concept, along with Third World demands for a "new international economic order", should be dismissed as socialist subversion.

In dismissing several key members of the US negotiating team, and initiating a complete review of the treaty in its draft form, the Reagan Administration is responding to complaints from groups such as the Council on Economics and National Security. Council officials were among the principal authors of a warning by the American Geological Institute to the two presidential candidates last September of a Soviet "resource war" aimed at depriving the United States of materials indispensable to its national defence.

Immediately after the election, the Reagan Administration set up a task force on national strategic minerals policy, which recommended a complete review of the draft Law of the Sea Treaty. One of the principal opponents of the treaty in its present form is a Washington attorney, Leigh Ratiner, acting head of the Department of the Interior's office of oceans and minerals under President Ford, and later the chief lobbyist on the treaty for the Kennecott Corporation.

Mr Ratiner is acting as counsel to the L-5 society, a 4,000-strong group of scientists and "futurists". The society is largely credited with stalling any effort by the Senate to ratify a draft treaty covering mining operations on the Moon, agreed in principle by the United Nations last year.

Particularly strong opposition to the treaty has been voiced by the US Steel Corporation of Pittsburgh, a leading partner in the Ocean Mining Associates consortium. According to the chairman of US Steel, William R. Roche, policies pursued by the Carter Administration tended to increase the country's strategic mineral vulnerability.

Attempts to scuttle the Law of the Sea negotiations have been fiercely attacked by Mr Carter's chief negotiator Mr Eliot Richardson, ex-ambassador to the United Kingdom, who warned that many multinational companies would face severe consequences if they tried to disrupt

agreements that had already been reached by US negotiators in good faith.

Many conservatives, however, see Mr Richardson's position as symptomatic of a "soft" approach to the Third World that does little to protect US economic and strategic interests. A change in strategy on the Law of the Sea Conference could also have implications for other negotiations.

Following the Antarctic Treaty of 1959, the 12 (now 14) nations laying claim to Antarctic territory have agreed voluntarily to restrain mining activities until an appropriate international treaty can be worked out. Yet this too is now being challenged by conservatives within the Reagan Administration, largely on the grounds that negotiations could result in the United States gaining less than it might expect from an international free-for-all, given its superiority in mining technology.

Delegates attending the Law of the Sea negotiations session in New York, although dismayed with the Reagan Administration's position, have agreed to

discuss the details of the draft treaty. Some had hoped that the Kennecott/Sohio deal, together with the influence of other international groups such as Royal Dutch Shell, might help to dilute the pressure of more nationally-oriented companies on the Reagan Administration.

Meanwhile, a group of lawyers, scientists, environmentalists and economists has issued a report calling for research into the character of the marine environment of the deep ocean, and its susceptibility to mining operations. The group points out that no explicit mention is made of the need for prior environmental assessment either in a bill on deep-sea mining rights passed recently in West Germany, or in draft bills now before French and British legislatures. It was also concerned that such nations might get together with the United States and other countries to set up their own institution for regulating mining activities before the Law of the Sea negotiations finally break down.

David Dickson

India mines the sea

New Delhi

India has become the first developing country to collect polymetallic nodules from the ocean floor, a distinction achieved by India's first research ship, the *Gaveshani*. The 1,900-tonne vessel, once a hopper barge and converted into a research ship in 1975, is operated by the National Institute of Oceanography (NIO) in Goa, one of the 40 institutes under the aegis of the Council of Scientific and Industrial Research.

NIO has been preparing the *Gaveshani* for nodule hunting since 1978 and the catch was made in January during her cruise from Goa to Mauritius, which was led by Dr S.Z. Qasim, director of NIO.

According to Dr Qasim, the *Gaveshani* located a nodule-rich region in the Indian Ocean about 700 miles from the mainland and reaching almost as far as Mauritius. During the return journey the ship collected more nodules from several locations at depths of 3,600–5,400 metres.

Besides nickel, cobalt, copper, manganese and iron, the samples are believed to contain traces of gold, but their precise metallic contents are now being analysed by eight Indian laboratories.

The *Gaveshani*'s nodule collecting gear

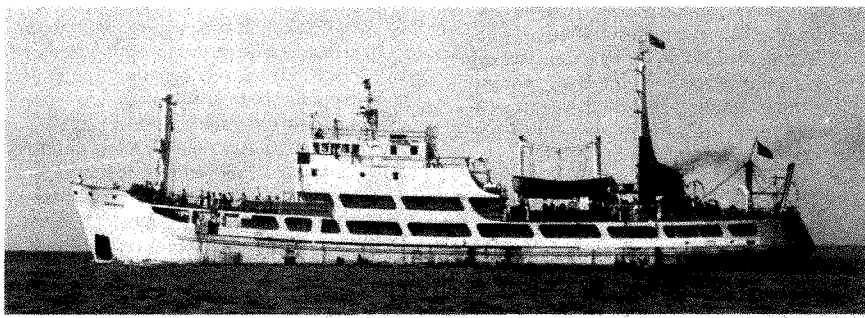
consisted of a 10,000-metre winch, underwater camera, dredges and some twenty "boomerang grabs", which go down to pick up the nodules and come up automatically. According to Dr Qasim, the surveyed seabed area had a nodule concentration of 1 to 5 kilogrammes per square metre, which is potentially economic for mining.

India is not ready to start mining proper, but NIO plans to map the nodule-rich region to assess its total mineral wealth. The operation should be speeded up next year when the *Gaveshani* is joined by a second research ship now under construction in West Germany. Since India imports all its nickel and cobalt and more than 60 per cent of its copper requirements, the discovery of a nodule basin in its close vicinity should be welcome news.

There is a long haul between picking up a few seabed nodules and being able to mine them commercially. Yet the *Gaveshani*'s demonstration has raised some eyebrows among nations that at present have a monopoly on ocean mining. It is likely that United Nations talks on the Law of the Sea will acquire a new dimension as a result of India's unexpected entry into the ocean mining club.

K.S. Jayaraman

Gaveshani — hunting for nodules



CORRESPONDENCE

Licking Darwin

SIR — How absurd to hold that Darwinism cannot be falsified. All it would take is a single clear demonstration of the *de novo* creation of a complex organism, well adapted to its environment, and the Darwinists would be legless.

G.P. VINSON

Department of Biochemistry and Chemistry,
The Medical College of St Bartholomew's
Hospital, London EC1, UK

Weed control

SIR — As May emphasizes¹, all possible steps should be taken to prevent the intrusion of noxious weeds into new regions. That article paid rather less adequate attention to weeds that are already established. Consider two hypothetical situations. Suppose it was discovered that one of these weeds contained a valuable fibre. Within a few years we would have a conservation programme to save the weed from extinction. Suppose enterprising adolescents found that one of the weeds contained a hallucinogen. The \$15 million quoted as the annual cost of water hyacinth control in Florida and Louisiana would be "chicken feed" compared with what drug control authorities would spend on elimination.

Water hyacinth has been studied for so long that components as dramatic as these are unlikely to be found in it. Nevertheless, it is a prodigious producer of biomass and is often conveniently located on broad areas of water from which mechanical collection is easy. Twenty one years ago² I made "... a plea for some reorientation of thinking in areas suffering from, or threatened with, water hyacinth. There is a case for directing some research towards finding uses for the pest instead of concentrating all efforts on its eradication."

Since then, the pest has made great headway. Less headway has been made on finding uses for it. Interest is, however, increasing. Chibbar and Singh³ fed cows on silage made from water hyacinth and paddy straw, and there have been several recent papers in the *Indian Journal of Animal Science* extending that work. Hyacinth is deliberately cultivated in China for pig feed. It is a component of the mixtures used for biogas fermentation in many countries. These and other uses are surveyed in *Making Aquatic Weeds Useful* (US National Academy of Sciences, 1976) and in *Handbook of Utilization of Aquatic Plants* (Food and Agricultural Organization, 1979). This is encouraging. But the amount of effort put into finding uses for water hyacinth and other weeds is a tiny fraction of the effort put into killing them. I do not suggest deliberate encouragement, only the assignment of comparable importance to both lines of approach.

N.W. PIRIE

Harpenden, Herts, UK

Cafeteria feeding

SIR — I wonder if I am alone in my dislike of the term "cafeteria feeding" used first, I think, by Rothwell and Stock in 1979 (Rothwell, H.J. & Stock, M.J. *Nature* 281, 31; 1979). According to the *Concise Oxford Dictionary* a cafeteria is a "restaurant in which customers fetch food and drink from a counter" which is hardly an accurate description of the behaviour of rats allowed to make their own choice of palatable items of food. I would like to suggest that, before it is too late, editors of scientific journals ask authors to use a more accurate description, for example to replace "cafeteria diet" by "a self-selected palatable diet" (abbreviated SSPD) or other equivalent term.

A.S. COLE

University of Bristol, UK

Cancer causation

SIR — The recent series of articles¹⁻³ on cancer causation prompts me to respond. Without defending Epstein and Swartz, I would like to offer several alternative explanations to some of the examples which Cairns used to argue his case that somatic mutations play little or no part in human carcinogenesis.

Although Cairns' novel hypothesis is to be welcomed as a stimulus to new experimental approaches, I feel that he has summarily dismissed a significant amount of evidence consistent with the somatic mutations theory of cancer. Furthermore, carcinogenesis is a multi-step process and there are probably several distinct molecular mechanisms that can "cause" and influence carcinogenesis, not the least of which are somatic mutations⁴.

Specifically, Cairns argues that the *apparent* lack of common fatal cancers in xeroderma pigmentosum (XP) individuals (the cells from whom lack proper DNA repair causing them to be hypermutable to many mutagens, including the ultraviolet component of sunlight) is evidence that the production of mutations in XP cells is probably irrelevant to carcinogenesis. There are at least two reasonable alternative explanations why XP patients rarely get cancers other than skin tumours⁵. The XP patients normally do not live long enough to manifest these internal tumours (an explanation cited but dismissed by Cairns) and most cancers are the result of initiation (mutagenesis) and promotion (amplification of an initiated cell into a clone where additional genetic or epigenetic changes can occur). One can explain the production of skin tumours in XP on this hypothesis since, using this concept, repaired sunlight-induced lesions in the skin cells of XP patients will lead to both cell killing and mutated viable cells.

In effect, the large numbers of killed cells will "force" the surviving, but initiated, cells to repopulate the necrotic tissue, giving rise to clones of dysfunctional cells. During that repopulation phase, some of these initiated cells are further exposed to more mutagens. Thus, the cytotoxicity caused by the inability to repair large amounts of sunlight-induced DNA damage can be considered an "indirect promoter", much in the manner of a partial hepatectomy⁶. On the other hand, internal cells of XP patients, which are exposed to

chemical mutagens that mimic sunlight-induced DNA damage, probably are not exposed to high enough levels constantly to cause significant cell killing, as occurs in the skin. In effect, these initiated cells will not be promoted! Only after long periods of time or exposure to chronic effective levels of exogenous or endogenous tumour promoters will internal tumours appear.

A second argument used by Cairns to support this thesis of genetic transposition was the discussion on Bloom's syndrome. I agree with Cairns that Bloom's syndrome might be a "better" model for studying carcinogenesis, but for a different reason. At the time of writing his paper, Cairns was not aware of the observation that Bloom's syndrome seems to be a "mutator" syndrome⁶, in that it spontaneously generates mutations in their somatic tissue. In effect, Bloom's cells can generate mutations in any proliferation cell of the body, with or without exposure to "environmental" mutagens. This would explain the wide range of tumours (and other clinical features) associated with the syndrome, and would also fit the initiation/promotion model, since each time a spontaneous mutant cell arises and then divides, it can enhance the probability of additional genetic imbalances.

JAMES E. TROSKO

Michigan State University,
East Lansing, Michigan, USA

1. Epstein, S.S. & Swartz, J.B. *Nature* 289, 127-130 (1981).
2. Cairns, J. *Nature* 289, 353-357 (1981).
3. Editorial *Nature* 289, 431-432 (1981).
4. Trosko, J.E. & Chang, C.C. in *Chemical Carcinogens and DNA* Vol. 2 (ed. Grover, P.L.) 181-200 (CRC Press, Boca Raton, Florida, 1979).
5. Trosko, J.E. & Chang, C.C. *Med. Hypothesis* 6, 455-468 (1980).
6. Warren, S.T. et al. *Proc. natn. Acad. Sci. USA*. (in the press).

Genes and free will

SIR — Asked to dissociate themselves from the neo-Nazi use of their names and theories, John Maynard Smith¹ and Richard Dawkins² have both done so. Professor Maynard Smith's unequivocal couple of sentences will be no surprise to scientists who know him as a man of principle and integrity; his letter will give no comfort to the National Front or any similar grouping.

It is a pity, though, that Dr Dawkins appears to think that the issue I sought to raise in my letter is of a parallel with the relatively harmless debate on cladistics and dialectics which has kept us entertained over recent months. If it were not for that dreary little edge of "reds under the bed" of which Stephen Jay Gould reminds us³, I suspect we could look forward to *Nature* with the same enthusiasm as once we had for *Beano*. However, the issues are not the same. Racism directly haunts the streets of our cities. In our laboratories neither Dawkins nor I live in the daily hazard that the obscene rubbish of the *New Nation* and its followers fosters in the East End or New Cross. When Dawkins disdains the ephemera of mere "human politics", or accuses his critics of being as guilty as the Nazis of dragging the elegant purity of his neo-Darwinism into the litter of the city streets, he does so at the peril of a repetition of the tragedies of the 1930s.

Of course we must all beware of the fallacy

Continued on page 432

1. May, R.M. *Nature* 290, 85-86 (1981).

2. Pirie, N.W. *Nature* 185, 116 (1960).

3. Chibbar, S.S. & Singh, G.D. *Indian Farming* 20, 24 (1971).

NEWS AND VIEWS

Nerves, myelin and multiple sclerosis

from Miranda Robertson

At the beginning of December, a select group of clinicians came together with a similar group of research biologists in Berlin to plumb the depths of their respective ignorance about multiple sclerosis and the glial cells of the central nervous system. Multiple sclerosis is baffling in almost every respect — its cause is unknown, its pattern erratic and its course extraordinarily variable. The target of the disease is the myelin that insulates the nerve fibres of the central nervous system, or the oligodendrocytes that produce it. The symptoms are thus due to blocked or unreliable conduction along nerves, but it is not known what induces the attack on the myelin; nor is it understood why, rather than steadily worsening, the symptoms should alternately flare up and partly subside over the progressive course of the disease; nor why some patients may remain free of symptoms for years after a first attack, whereas others will suffer repeated relapses and die within a decade.

Two of the principal aims of the meeting*, whose main topic was the relationship between neurones and glia but which also included much discussion on viral and immunological mechanisms, were to try to establish, first, what is going on in multiple sclerosis patients, and second, what might be done about it.

Probably the most cogent view of the characteristics of the disease is due to the immunologists, whose contribution has recently been discussed in *Nature* by Bloom¹. The pattern of exacerbation and remission of symptoms, for example, may represent defective (and thus irregularly oscillating) regulation of an immune response, manifesting itself as recurring episodes of inflammation — an explanation consistent with the efficacy of steroids during acute attacks (B. Arnason, Chicago University).

The cause of the inflammation and tissue

destruction could be a latent virus infection and the immune response to it, or an autoimmune attack on the central myelin that might or might not have been stimulated in the first instance by viral infection. In animals, the pathological changes can be mimicked by some viral diseases, or by immunization with central myelin — though in the latter case, the disease follows a chronic relapsing course only in animals with genetic immune deficiencies. (This may be significant in the light of evidence in man for linkage with specific HLA antigens: see below.)

However, multiple sclerosis itself is not associated with a specific immune response either to a particular virus or to a component of the central nervous system; nor is there any evidence that immunosuppressive treatment can alleviate acute attacks of the disease. This raises two questions. The first is that of the relevance of any of the existing animal models to the human disease. The second is whether there is any point in looking for immune correlates of multiple sclerosis without having identified a specific antigenic target: immune activity in itself is a very poor guide to aetiology, since tissue degeneration from any cause (including, for example, stroke) will attract the attention of the immune system.

Nonetheless, many laboratories have looked for changes in the immune response that may be associated with multiple sclerosis, and one of the most recent fruits of such research has been the discovery that exacerbations of the disease coincide with a reduction in the level of circulating immunoregulatory T cells². This may mean that the failure of immune regulation unleashes the attacks (in which case, however, why specific to central myelin?), or alternatively that suppressor T cells and oligodendrocytes share an antigen and the disappearance of the immunoregulatory lymphocytes is a concomitant and not a cause of the attack. Many participants at the meeting took the view that the latter was the more interesting possibility

because it could lead to the identification of the antigenic target — an advance whose importance it would be hard to exaggerate.

There is, of course, a strong possibility that there is no single cause of multiple sclerosis. It was generally agreed, for example, that if the disease is a response to viral infection it is very likely that more than one virus can elicit it. But some very preliminary genetic evidence suggests that the heterogeneity of multiple sclerosis may run deeper than that. Nowadays, if nothing else is known about a disease — in fact, particularly if nothing else is known — its HLA association is. Multiple sclerosis in Europe is in linkage disequilibrium with HLA antigens D2 and D4 (ref. 3), but more recent investigations have revealed a further link, with DR3. The significance of these associations is unknown, but their weakness, together with genetic evidence from family studies, and in particular from twins, strongly suggest that HLA D2 and D4 are in linkage disequilibrium with some other, unknown genetic factor predisposing to the disease (W.I. McDonald, London University). The case of DW3, on the other hand, may be different: DW3 is associated with two known autoimmune diseases, myasthenia gravis and autoimmune thyroiditis; there is thus a possibility that its prevalence reflects a defect in immune regulation predisposing to autoimmune attack — the exact target presumably depending on other factors.

If the ultimate cause of multiple sclerosis remains unknown, can anything be done in the meantime about its proximal cause? Can conduction be restored in demyelinated nerves, or dead oligodendrocytes be replaced? Drugs that enhance conduction have been tried, but without much success. One of the most recent efforts in that direction has been with 4-aminopyridine, which blocks potassium channels and prolongs the action potential⁴. This facilitates

Miranda Robertson is an associate editor of *Nature*.

*The Dahlem Conference on 'Neuronal-Glial Cell Interrelationships' was held in Berlin on November 30–December 5 1980. A Dahlem Workshop Report on the conference will be published.

conduction but unfortunately at the same time promotes seizures. It is, in any case, doubtful how far simply overcoming conduction block could restore useful function, since both the pattern and the spread of impulses would inevitably be grossly abnormal.

Some natural short-term recovery of function, on the other hand, is to be expected on the basis of recent investigations on the distribution of sodium channels in myelinated nerves. In the peripheral nerve, sodium channels are clustered at the nodes of Ranvier, being scarce or absent in the membrane under the myelin sheath⁵. This means that immediately after the demyelination of a peripheral nerve, conduction is completely blocked and it is not until some days afterwards that sodium channels begin to appear at the internode region and action potentials can be propagated across it (T.A. Sears, London University). If the same differential distribution of channels holds in the central nervous system (and there is some evidence that it does⁶), the exacerbating and remitting pattern of multiple sclerosis may partly reflect the

length of time it takes for sodium channels to redistribute themselves after demyelination.

Real recovery could be brought about only by the proliferation of oligodendrocytes and remyelination of axons. Although Schwann cells are known to proliferate and remyelinate peripheral fibres, and central remyelination commonly occurs in other demyelinating diseases, it does not seem to occur in multiple sclerosis. This suggests that multiple sclerosis may destroy the oligodendrocytes which, unlike Schwann cells, cannot be replaced. This reflects the more advanced state of differentiation of the oligodendrocyte. Not only can Schwann cells in culture be induced to proliferate, both by contact with axonal membrane (R.P. Bunge, Washington University) and by a soluble pituitary protein⁷, but they can also be induced by co-culture with nerves to increase 1,000-fold the synthesis of the principal myelin protein; or, by co-culture with muscle cells, to synthesize acetylcholine. Cultured oligodendrocytes, on the other hand, will synthesize the principal central myelin proteins in the absence of

neurones, although they do seem to depend on astrocytes to support proliferation (R. Mirsky, London University). What signals either Schwann cells or oligodendrocytes *in vivo* is quite unknown, although with the recent development of specific markers for central and peripheral neurones and glia, this question is now open to investigation *in vitro*.

Whether such investigations will make any direct contribution to the treatment of multiple sclerosis, or whether the capacity for recovery in the central nervous system lies in its inherent plasticity rather than in regeneration, only more research can show. In the meantime, everyone agreed that prevention is always preferable to cure; in the case of multiple sclerosis, it may be the easier to achieve. □

1. Bloom, B. *Nature* **287**, 275 (1980).
2. Antel, J.P., Richman, D.P., Medof, W.E. & Arnason, B.G.W. *Neurology* **28**, 106 (1978).
3. Batchelor, J.R., Compston, A. & McDonald, W.I. *Br. med. Bull.* **34**, 279 (1978).
4. Sherratt, R.M., Bostock, H. & Sears, T.A. *Nature* **283**, 570 (1980).
5. Chiu, S.Y. & Ritchie, J.M. *Nature* **284**, 170 (1980).
6. Kocsis, J.D. & Waxman, S.G. *Nature* **287**, 349 (1980).
7. Brookes, J.P., Lemke, G.E. & Balzer, D.R. Jr *J. biol. Chem.* **255**, 8374 (1980).

A new model for phototransduction

from Peter G. Lillywhite

IN this issue of *Nature*, R. Payne and J. Howard (see page 415) put forward a new model for phototransduction which provides an elegant alternative to the cascade model of Hodgkin¹. A successful model for phototransduction in vertebrates and invertebrates needs to explain (1) the latency of the response, (2) the increase in sensitivity to light steps and flashes with dark adaptation, and (3) that large increases in sensitivity can occur with only small changes of time scale. In the model invented by Hodgkin and fitted to *Limulus* receptors in collaboration with Fuortes¹, a cascade of stages, each occurring with an exponential delay, was assumed. Several stages are required because a single step would not, for example, exhibit latency. It turned out that about ten exponential stages were required to mimic the responses of *Limulus* receptors. Strong illumination was assumed to decrease the time constants of most, but not all of the exponential stages of transduction, causing the decrease in sensitivity observed with light adaptation.

It was immediately attractive to give physical reality to the hypothetical chain of exponential stages or 'events' as chemical reactions. Borsellino *et al.* showed in 1965 that, if it is assumed that the time constants of the exponential delays represent the rates of chemical reactions, the effect of temperature upon voltage response fits the classical Arrhenius relationship². Although they took the necessary precaution of pointing out that consistency

with the Arrhenius relation was not proof that the hypothetical chain of chemical reactions actually existed, chemical reaction chain models are nevertheless highly popular. For example, Baylor *et al.* recently (1974) modelled the responses of cones in the turtle retina using a chemical reaction cascade model³.

Payne and Howard now propose an alternative mathematical and conceptual model to that of Fuortes and Hodgkin and other similar cascade models. The model is attractive, because it describes locus photoreceptor responses (which are very similar to those of *Limulus* modelled by Fuortes and Hodgkin) almost perfectly, accounting for the three characteristics listed above, but does not require a chain or cascade of many hypothetical stages. The cascade models of Fuortes and Hodgkin and Baylor *et al.* have an unattractive side to them, which is the large number of implied transduction steps or reactions. To fit the photoreceptor responses they studied, Fuortes and Hodgkin required about ten steps and Baylor *et al.* required nine (six for creation and three for destruction of ion channel blocking particles). When Payne and Howard tried to use these models to fit locust photoreceptor responses, between 13 and 20 (Fuortes and Hodgkin model) and between 20 and 48 (Baylor *et al.*'s model) steps were required — a seemingly implausible number. Payne and Howard's

model is simple and minimal in its complexity and assumptions: photoreceptor impulse responses are assumed to be logarithmically transformed Gaussians. The only variables are s , the sensitivity, which is fixed for a given cell, and μ and σ , the mean and standard deviation of the transformed Gaussian. Furthermore, the effect of light adaptation and temperature change can be modelled by changing a single parameter, μ , which changes the time scale of the response and could represent the rate constant of a single transduction process.

Mechanistically, as well as mathematically, the Payne and Howard model is a simple working hypothesis which fits available data and is strikingly similar to recent suggestions by Hamdorf and Kirschfeld⁴ arising from their experiments on fly photoreceptors. Payne and Howard assume that absorption of a photon by a rhodopsin molecule leads to release of a chemical transmitter, the concentration of which smoothly rises to a peak before declining. The transmitter molecules open the ion channels responsible for the change in membrane potential during illumination. In insects and those other invertebrates which have rhabdomeric photoreceptors, illumination causes depolarization. However, in order to adapt this model to vertebrate ciliary photoreceptors, which hyperpolarize during illumination, one need only assume that the transmitter molecules close, rather than open ion channels. The threshold transmitter concentration for activation varies

Peter G. Lillywhite is Lecturer in Zoology at Sheffield University.

from one ion channel to another according, it is assumed, to a gaussian distribution. This variability is responsible, according to the Payne and Howard model, for the long-lasting nature of the membrane potential produced by an instantaneous light pulse. The long latency of phototransduction, which can be tens of milliseconds for small responses in dark-adapted receptors, is assumed to be due to the time required for transmitter concentration to exceed the activation thresholds of the ion channels. Hamdorf and Kirschfeld recently pointed out that the rapidity of micro-spectrophotometrically observable pigment conversion (around 0.5 ms) and diffusion cannot account for the long latency of phototransduction in flies. In showing that simultaneous absorption of two photons by the same rhabdomere microvillus causes a discrete reduction in latency, Hamdorf and Kirschfeld concluded that each rhabdomere microvillus may represent a single functional unit, which generates transmitter molecules by means of a single enzymatically activated transduction step⁴. As in the Howard and Payne model, Hamdorf and Kirschfeld attribute the long latency to the time taken for the concentration of transmitter generated to exceed a threshold level.

An additional requirement for any realistic model of transduction in insects is that the first 3–5 ms of transduction latency consist of reversible events localized at the level of the photopigment molecule, since Hamdorf and Kirschfeld have recently shown that reconversion of metarhodopsin to rhodopsin within this period prevents the rhodopsin to metarhodopsin conversion having any effect⁵. Thus, no irreversible events, such as release of 'live' transmitter, occur in this period. It is attractive to assume that the remaining few milliseconds before onset of the response are indeed due to a single enzymatically activated generation of transmitter, as suggested by Hamdorf and Kirschfeld⁴ and as fitted by Payne and Howard's mathematical model. No transduction enzyme has yet been fully characterized, although there is some evidence that a specific protein is involved in the intermediate steps of phototransduction⁶.

It would be interesting to know how well the Payne and Howard model fits the responses of other invertebrate photoreceptors and whether it can be used to model the responses of vertebrate photoreceptors, because it predicts a much simpler model of transduction than the reaction chain hypotheses. □

1. Fuortes, M.G.F. & Hodgkin, A.L. *J. Physiol., Lond.* **172**, 239 (1964).
2. Borsellino, A. & Fuortes, M.G.F., Smith, T.G. *Cold Spring Harb. Symp. quant. Biol.* **30**, 429 (1965).
3. Baylor, D.A. & Hodgkin, A.L., Lamb, T.D. *J. Physiol., Lond.* **242**, 685 (1974).
4. Hamdorf, K. & Kirschfeld, K. *Z. Naturforsch.* **35c**, 173 (1980).
5. Hamdorf, K. & Kirschfeld, K. *Nature* **283**, 859 (1980).
6. Pāk, W.L. & Ostroy, S.E., Deland, M.C., Wu, C.-F. *Science* **194**, 956 (1976); Fung et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 152 (1981).

Second sound in ³He

from P. V. E. McClintock

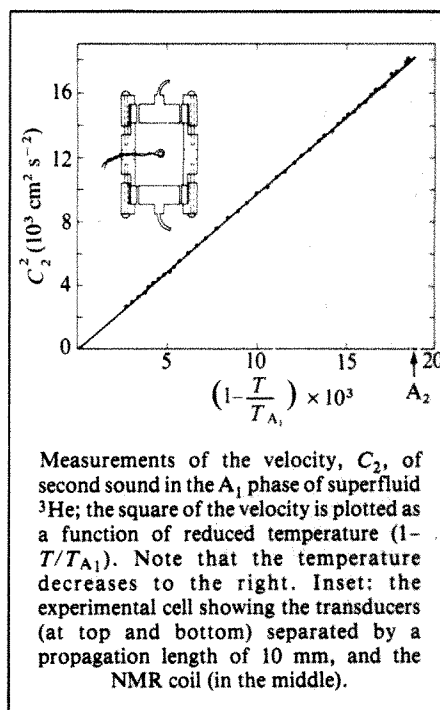
SECOND SOUND is an unusual type of propagating wave mode, which can occur in superfluids, involving fluctuations in the local temperature and entropy of the medium rather than in the local density and pressure as found in a conventional sound wave. The phenomenon has been investigated, both experimentally and theoretically, in the common helium isotope, ⁴He, over many years; but its observation in ³He, whose ultra-low temperature superfluidity is still a relatively recent (1972) discovery, was reported for the first time by Corruccini and Osheroff only last year (*Phys. Rev. Lett.* **45**, 2029).

At temperatures not far below their superfluid transitions (~ 2 K for ⁴He, 2 mK for ³He), both liquid heliums behave much as though they were composed of two separate but freely interpenetrating liquids: a superfluid component (anisotropic in the case of ³He) with no viscosity or entropy and a normal fluid component possessing the properties of a (fairly) normal liquid. This type of two-fluid system is able, in principle, to support two quite distinct wave modes. Where the two components vibrate in unison with each other, the result is an ordinary density wave known as 'first sound'. If, on the other hand, the two components move in antiphase, with the overall liquid density remaining essentially constant, the characteristic wave propagation velocity turns out to be quite different, corresponding to 'second sound'.

Although second sound is very readily generated and detected in superfluid ⁴He, calculations had suggested that the corresponding mode in the principal (A and B) phases of superfluid ³He would travel exceedingly slowly and be very strongly attenuated as it travelled through the liquid. Consequently, although second sound should, in principle, exist in ³He, it would quite probably prove impossible to detect it in practice. Liu predicted (*Phys. Rev. Lett.* **43**, 1740; 1979), however, that within a very narrow region of the phase diagram, the A₁ phase, the liquid might, in fact, be capable of supporting a freely propagating second sound wave — and it would be second sound of an entirely novel and particularly interesting kind.

The A₁ phase is formed under the influence of a strong magnetic field as an intermediate stage in the transition between normal liquid ³He and the superfluid A phase. The superfluid component of the latter can itself be considered as being made up of two sub-components in which the constituent pairs of ³He atoms lie with their nuclear magnetic moments lying either parallel or anti-parallel to an applied magnetic field. These 'up-spin' and 'down-

spin' superfluids seem, in many respects, to behave almost independently of each other. When normal liquid ³He is cooled in a magnetic field, the superfluid transition to the A phase is found to take place in two stages, at temperatures T_{A1} and T_{A2} whose separation is proportional to the applied field, and this is believed to correspond to the two superfluid sub-components having slightly different transition temperatures. Thus, for temperatures in the narrow range (typically 50 μ K) between T_{A1} and T_{A2} , only one of the spin species will be superfluid — this is known as the A₁ phase.



Liu pointed out that second sound in the A₁ phase, if it occurred, would carry an associated magnetic disturbance: counterflow of the normal and superfluid components would necessarily involve a spin density wave because only one preferentially aligned spin species would be present. He was able to predict that, for a hybrid second sound wave/spin density wave of this nature, the velocity would be much higher and the attenuation much smaller than would otherwise be anticipated for second sound in ³He.

It is this hybrid wave mode whose existence Corruccini and Osheroff have now successfully demonstrated at the Bell Telephone Laboratories, Murray Hill. They generated the mode by vibrating a capacitatively coupled porous membrane in the liquid. The precise mechanism by which this type of transducer operates is still somewhat ambiguous but, if one assumes that the membrane couples only to the normal fluid component, it is then clear that any motion of the membrane will result in a relative velocity of the normal

P. V. E. McClintock is in the Physics Department at the University of Lancaster, UK.

and superfluid components. The primary detection system depended on the same principles, but operating in reverse.

The received signals were weaker than anticipated, but yielded velocities which, as shown in the figure, appear to be in excellent agreement with the prediction, based on the theory of Liu, that the square of the velocity should be directly proportional to the reduced temperature ($1 - T/T_A$). Some highly significant additional information was obtained from a small NMR coil placed in the middle of the experimental cell, which enabled the authors to monitor the magnetic disturbance which occurred as the second sound wave passed by on its way from the generator to the detector. On the basis of

careful analysis of the relative shapes of the NMR and acoustic signals they tentatively conclude that, quite contrary to expectation (Levin *Phys. Rev. Lett.* **34**, 1002; 1975), the paired spins forming the superfluid in the A_1 phase are in reality aligned 'up' along the magnetic field, rather than 'down'. This conclusion should be treated with caution until corroborated by some other means, but it is certainly intriguing.

Although several features of the experimental results still await detailed interpretation, there now seems no doubt that Liu's exotic wave mode exists in reality and that the work will lead to a significantly enhanced understanding of superfluid ^3He . $\square$

Ice core studies: dating the past to find the future

from W. Dansgaard

How did atmospheric carbon dioxide influence the climate in the past? Will the two large ice sheets, in Antarctica and Greenland, be stable enough to withstand a possible climatic warming? If not, how fast could they disintegrate? How old are they? How abrupt was the termination of the last interglacial, the Eem? When did it happen, and why? Did volcanic activity trigger that event? Is the Earth ripe for another catastrophic turn into glacial conditions? How did cosmic radiation change in the past? What were pre-industrial deposition rates of the various chemical elements that are now being emitted by modern industry?

These, and many other important problems may be partially, if not completely solved by investigation of the more than 3,000 m thick ice sheets which are proving richer sources of information about past atmospheric conditions than any other sequence of Quaternary sediments on Earth. All kinds of fall-out from the atmosphere, such as air-borne continental dust and biological material, volcanic debris, sea salts, cosmic particles and cosmic ray-produced isotopes, are deposited on the surface along with the snow and each year a new layer is added to the old. The snow pack is gradually compressed into solid ice with small cavities containing samples of atmospheric air. In the coldest areas of the ice sheets, the impurities remain in the ice as deep-frozen indicators of the chemical composition and the physical condition of the atmosphere at the time of deposition. Nothing is added, nothing runs off or is displaced, and no chemical reaction takes place; in fact, the composition of the ice layers changes only

by decay of the radioactive impurities and by extremely slow diffusion processes in the ice crystal lattice.

The ice layers sink into the ice sheet in an undisturbed sequence with continuous horizontal stretching and consequent thinning. In areas with no melting at the bedrock, the layers approach zero thickness close to the bottom. This is the reason why an ice core drilled through the coldest part of Greenland ice sheet makes it possible to establish continuous and detailed time series of many geophysical-chemical parameters, reaching several hundred thousand years back in time: CO_2 content in the atmosphere^{1,2}; climatic changes in terms of surface temperature³⁻⁵ and accumulation variations⁶; chemical composition of the atmosphere⁷; volcanic activity and its cooling effect in the troposphere⁸ and cosmic radiation flux¹. In central East Antarctica the oldest ice may even be an order of magnitude older than in Greenland, but the snow accumulation rate is so low⁹ — in some areas even zero or negative — that an Antarctic ice core does not necessarily represent a continuous sequence of deposition.

Dating of an ice core is prerequisite for proper interpretation of a data series. A number of dating methods have been developed on the basis of ice flow models, and analyses of isotopes, volcanic debris and continental dust in ice cores¹⁰. Most accurate are the stratigraphical methods, that is, detection and counting of the individual annual layers downwards from the surface. In areas with high accumulation ($> 25 \text{ cm ice yr}^{-1}$) and no melting, these methods allow absolute dating to within a few years per thousand if the ice core is in perfect physical condition. This

accuracy is required when high-frequency components of the time series are to be compared with similar features in other time series. For example, the acidity of annual layers could be reconciled with historical records of volcanic eruptions within the past millennium⁸ only because each layer was dated back to near the correct year of deposition¹⁰. So far, no ice core has been absolute-dated beyond AD 553.

Under favourable conditions the annual stratification will probably be detectable far beyond 10,000 yr BP, but, at present, ice from the early stages of the last glaciation must be dated by radioisotopes. However, the conventional low-level β -counting technique requires preparation of much larger amounts of ice than is available in an ice core¹¹. It is therefore a great step forwards that the limit of radioisotope detection has been lowered by orders of magnitude in recent years by means of an accelerator/mass spectrometer combination that identifies and detects the radioactive atoms themselves instead of the β -particles emitted when they decay¹²⁻¹⁵.

This makes it possible to perform ^{14}C dating back to 50,000 yr BP using as little as $1 \text{ cm}^3 \text{ CO}_2$ which is the approximate CO_2 content in 5 m of a 10 cm ice core. This core length may span only a few hundred annual layers at depths corresponding to 50,000 yr of age in the central parts of the large ice sheets¹⁶. Ice from the early stages of the last glaciation has to be dated by longer-lived isotopes, of which ^{36}Cl is the most promising. As the half life is relatively favourable (310,000 yr), a ± 3 per cent detection accuracy is within reach¹⁵; and an 0.5 m core increment is sufficient. Yet, the corresponding dating uncertainty will be as high as $\pm 13,000$ yr, even if the production rate of ^{36}Cl in the atmosphere is assumed to have remained constant. If the production rate at the time of deposition were 10 per cent different from the present value, it implies a dating error of no less than 43,000 yr.

This difficulty may be partly overcome if the ^{10}Be (half life 1.5×10^6 yr) concentration is measured in the same ice. The ^{36}Cl to ^{10}Be concentration ratio may then be used as an age indicator that is essentially independent of the production rate, because both isotopes are produced in the atmosphere by impact of cosmic radiation, and both of them are scavenged by precipitation after a short residence time in the atmosphere¹. Nevertheless, when accounting for another 3 per cent uncertainty on the ^{10}Be analysis, we end up with a total dating uncertainty of $\pm 30,000$ yr on an individual ice sample. Owing to the small amount of ice needed, it may be feasible to carry out independent datings on, say, 10 ice core samples of approximately the same age around 100,000 yr, thereby reducing the statistical uncertainty on the mean value to $\pm 10,000$ yr, which can be extremely useful for checking the validity of ice flow models. A complete $^{36}\text{Cl}/^{10}\text{Be}$ profile along the lower part of an

W. Dansgaard is in the Geophysical Isotope Laboratory of the University of Copenhagen.

ice core, in which the entire transition (120,000 to 60,000 yr BP?) from the last and really warm interglacial (the Eem) to full glacial conditions is revealed by oxygen isotope ratio changes, would probably lead to the long-needed climatic counterpart to the continental ice volume records derived from analyses of foraminifera in deep-sea cores¹⁷⁻¹⁹.

These records show that at the end of the Eem some 120,000 yr ago the amount of continental ice began to increase, and that it grew from less than today to more than double the present amount within less than 10,000 yr. The drastic climatic deterioration, which triggered off the enormous glacier advance, may have taken place within a much shorter time interval²⁰, perhaps a century or less, to judge from a Greenland isotope record²¹ and from pollen records²²⁻²⁴. However, these records are poorly dated before 50,000 BP, and the possibility of a discontinuity in the layer sequences must be taken into account, the more so as the suggested duration (one millenium only) of the cold period that succeeded the Eem is inconsistent with the build-up of huge amounts of continental ice.

It is extremely important to study why and how abruptly the Eemian interglacial came to an end, because it might change our conception of atmospheric stability, and describe a possible mode of termination of the present interglacial. The new radioisotope ice dating technique will not be able to reveal the duration of the past climatic shift, but a counting of annual layers, deposited in a continuous sequence throughout the shift, may be feasible by modified detection techniques, because such layers, each 1 mm thick, probably exist some 80 m above the bottom of the ice sheet in central Greenland¹⁶. □

The origin of cosmic gamma-ray bursts — new Soviet results

From B. J. Teegarden

SINCE their accidental discovery in 1973 by Vela spacecraft, gamma-ray bursts have continued to be a mysterious and intriguing phenomenon. Lasting from a few seconds to a few tens of seconds, these bright flashes of gamma radiation occur randomly every one to two weeks. With one possible exception the identification of the sources of these events remains unsolved.

A multi-spacecraft network including experiments designed in the United States, France and the Soviet Union has been used to time the arrival of the gamma-ray burst wavefront to determine the arrival direction of the radiation. The only burst for which a source identification was made is the very unusual event of 5 March 1979 which displayed a very sharp, intense radiation spike lasting only 0.1 second and was followed by 8 second oscillations lasting for approximately 100 seconds. The arrival direction of this burst coincided with the remnant of a supernova explosion (labelled N49) in the Large Magellanic Cloud, our nearest neighbouring galaxy. This, however, was a highly unusual event, and its identification is still not universally accepted. At this time no other gamma-ray burst source identifications have been made. This is in spite of the fact that arrival directions for a number of events have been precisely determined.

A number of different models have been put forth to explain the energy source of these bursts. Two of these may be briefly summarized: the first hypothesizes that a solid body falls onto the surface of a neutron star providing a nearly instantaneous release of energy. The second proposes that infalling material from either a companion star or the interstellar medium gradually builds up a surface layer at the poles of a neutron star. Sufficient accumulation of material over a period of time leads to an unstable condition wherein a thermonuclear explosion can occur. In either case the resultant heating of the atmosphere immediately above the surface of the neutron star is sufficient to cause the release of an intense burst of gamma radiation.

Some dramatic new observations of gamma-ray bursts have recently been reported by E. P. Mazets and his colleagues at the Ioffe Physical-Technical Institute in Leningrad (see this issue of *Nature*, p.378) that shed important new light on the question of the origin of these events. Mazets and his co-workers have measured

the spectra of a large number of bursts using detectors on board the Soviet Venera 11 and 12 spacecraft. They have found remarkable emission and absorption features in 26 of the 80-90 events that were measured. Most of the spectra display pronounced absorption in the 30-70 keV range. This has been interpreted as being due to a phenomenon known as cyclotron absorption.

It is expected that extremely strong magnetic fields will surround a neutron star. This follows if the magnetic flux is conserved during the collapse of the parent star down to the neutron star. A simple extrapolation from the parent star to its compact offspring yields a magnetic field in the 10^{12} gauss range, a value far beyond any achieved in the laboratory. An electron in such a field will be constrained to spiral tightly about the magnetic field line. Quantum theory predicts that the electrons will be confined to discrete energy levels and that the transitions between these levels will be accompanied by either the emission or absorption of radiation. In a 5×10^{12} gauss magnetic field the characteristic energy of the transition is ~ 50 keV. The absorption and emission features in gamma-ray burst spectra observed by Mazets and his colleagues are therefore consistent with cyclotron processes in the strong magnetic fields expected to exist near the surface of a neutron star.

In addition to the 'cyclotron' emission and absorption features, Mazets and co-workers, as well as a group at Goddard Space Flight Center in the US, have seen in a smaller group of events (about 6) a broad emission line in the 400-450 keV range. When a positron and electron encounter each other and annihilate, they give off a characteristic emission at 511 keV known as annihilation radiation. Any such radiation originating near the surface of a neutron star will lose energy as it escapes from this intense gravitational field. This energy change is known as a 'gravitational redshift'. Present models predict that the magnitude of this shift will be 15-20% for a neutron star whose mass is the same as the Sun. Any 511 keV annihilation radiation produced near a neutron star surface would, therefore, emerge from the star's gravitational field with an energy in the 400-450 keV range.

It should be emphasized that these results are very recent and that the interpretations discussed here are not necessarily unique. These Soviet results, however, seem to point very strongly towards a neutron star origin for cosmic gamma-ray bursts. □

B. J. Teegarden is in the Laboratory for High Energy Astrophysics NASA/Goddard Space Flight Center.

- Oeschger, H. *Proc. Symp. chem. and radioact. Dating Am. Chem. Soc., Houston, Texas (March 1980)*.
- Delmas, R.J., Ascencio, J.-M. & Legrand, M. *Nature* **284**, 155 (1980).
- Johnsen, S.J., Dansgaard, W., Clausen, H.B. & Langway, C.C. *Nature* **235**, 429 (1972).
- Dansgaard, W. *et al. Nature* **255**, 24 (1975).
- Lorius, C., Merlivat, L., Jouzel, J. & Pourchet, M. *Nature* **280**, 644 (1979).
- Reeh, N. *et al. J. Glaciol.* **20**, 27 (1978).
- Herron, M.M., Langway, C.C., Weiss, H.V. & Cragin, J.H. *Geochim. cosmochim. Acta* **41**, 915 (1977).
- Hammer, C.U., Clausen, H.B. & Dansgaard, W. *Nature* **288**, 230 (1980).
- Barkov, N.I. *US Antarc. J.* **10**, No. 2, 55 (1975).
- Hammer, C.U. *et al. J. Glaciol.* **20**, 3 (1978).
- Oeschger, H., Stauffer, B., Brucher, P. & Moell, M. *J. Glaciol.* **17**, 117 (1976).
- Proc. 1st Conf. on Radiocarbon Dating with Accelerators* Univ. Rochester (ed. Gove, H.E.) (1978).
- Raisbeck, G.M., Yiou, F., Frunau, M., Lieuvain, M. & Loiseux, J.M. *Nature* **275**, 731 (1978).
- Bennett, C.L. *et al. Science* **201**, 345 (1978).
- Finkel, R.C., Nishiizumi, K., Elmore, D., Ferraro, R.D. & Gove, H.E. *Geophys. Res. Lett.* **7**, 983 (1980).
- Dansgaard, W., Johnsen, S.J., Clausen, H.B. & Gundestrup, N. *Meddelelser om Gronland* **197**, No. 2, 1 (1973).
- Emiliani, C. *J. Geol.* **74**, 109 (1966).
- Hays, J.D., Imbrie, J. & Shackleton, N.J. *Science* **194**, 1121 (1976).
- Duplessy, J.-C. in *Climatic Change* (ed. Gribbin, J.) 46 (Cambridge Univ. Press, 1978).
- Dansgaard, W. & Duplessy, J.-C. *Boreas* **10**, No. 2 (1981).
- Dansgaard, W., Johnsen, S.J., Clausen, H.B. & Langway, C.C. *Quat. Res.* **2**, 296 (1972).
- Andersen, S.T. *Medd. Dansk Geol. Foren.* **15**, 90 (1969).
- Wijmstra, T.A. *Acta bot. neerl.* **18**, 511 (1969).
- Woillard, G. *Quat. Res.* **9**, 1 (1978).

Plasmids and bacterial pathogens

from J.R. Saunders

THE way in which plasmids confer pathogenicity characteristics on host bacteria is of considerable concern in both developed and developing countries, and provided the central theme at a recent conference* in the Dominican Republic. Various aspects of plasmid biology were covered, ranging from the mechanisms of genetic transposition and plasmid replication to the epidemiology of antibiotic-resistant bacteria in human and animal populations.

Plasmids probably determine only a small proportion of all bacterial virulence traits. However, they may provide sufficient additional genetic information to convert relatively innocuous organisms into pathogens. This is particularly true of enteropathogenic strains of *Escherichia coli*, the prime cause of travellers' diarrhoea and other economically important diseases of humans and domesticated animals. The distressing symptoms of these conditions are largely produced by one or more plasmid-specified enterotoxins, which include a single type of heat-labile toxin (LT) and probably at least two distinct heat-stable toxins (ST). S. Falkow (Stanford) reported a striking similarity between the gene for LT and that for the cholera toxin, despite the fact that the former is always plasmid-borne in *E. coli* and the latter chromosomally determined in *Vibrio cholerae*. The difference in disease patterns, apparently specified by the same virulence determinant, may reflect the different methods by which the two causative agents actually deliver the toxin to their human hosts. This emphasizes how heavily the expression of pathogenicity determinants depends on the genetic and physiological infrastructure into which they become inserted. It is interesting in this context that certain characteristics, such as particular chromosomally determined somatic (O) antigen types, are found habitually among enterotoxigenic strains of *E. coli*.

Haemolytic strains of *E. coli* can frequently be isolated from infected body fluids. However, the exact role of haemolysins in the pathogenicity of these bacteria is not clear. In *E. coli*, α -haemolysins are determined either by the chromosome or by Hly plasmids. Such plasmids isolated from different strains have been shown by W. Goebel (Wurzburg) to share a common haemolysin determinant. This consists of three cistrons, *cisB* (1,500 base pairs) determining a protein located in the outer membrane and probably responsible for the transport of haemolysin, *cisA* (2,500 base pairs) encoding a precursor of α -haemolysin, and *cisC* (about 500 base pairs) which is apparently required for the

conversion of the *cisA* product to active α -haemolysin. Despite the wide distribution of the Hly determinant it does not seem to constitute part of a transposon (see De La Cruz *et al.* *J. Bact.* **143**, 825; 1980). This contrasts with one of the ST genes which is known to be transposable (So *et al.* *Nature* **277**, 453; 1979).

The ability of enteric bacteria to adhere to the gut wall is a primary requirement for their pathogenicity. The best studied of such plasmid-mediated determinants are those for the filamentous surface protein antigens K88 and K99, which enable enteropathogenic *E. coli* to adhere to the mucosa of the small intestine of pigs and calves respectively. Similar pilus-like structures, termed colonization factors (CFAI and II), are plasmid-encoded in human enteropathogenic *E. coli*. Two groups (J. van Embden, Balthoven and G. Dougan, Dublin) reported the cloning and genetic analysis of the K88 and K99 antigens. Dougan has found that at least four genes, organized into two operons, are involved in the expression of the K88 antigen in *E. coli* minicells. Operon I comprises three genes transcribed from *adhA* (encoding a protein of molecular weight 70,000 which may be the basal structure for membrane attachment of the antigen) through *adhB* (29,000-MW protein) to *adhC* (17,000-MW protein). *adhC* may encode a positive regulator for operon II containing a single gene *adhD* determining the 23,000-MW subunit of the K88 antigen itself.

In addition to adhering to the intestine, some enteric pathogens are able to invade the gut tissues. D. Kopecko (Washington) described plasmids of about 120 megadaltons in *Shigella sonnei*. Loss of these plasmids results in transition of host strains from the virulent, smooth colonial form I to a virulent, rough colonial form II. Such plasmids apparently determine the synthesis of the form I somatic (O) antigen required for invasion of colonic epithelium.

The potential of bacteria as invasive pathogens is obviously enhanced by their ability to withstand the antibacterial properties of serum. A number of unrelated antibiotic-resistance plasmids and temperate bacteriophages, such as λ , confer on host bacteria resistance to complement (activated by either classical or alternative pathways). K. Timmis (Berlin) showed that serum resistance conferred by the conjugative plasmid R6-5 is determined by the product of the *traT* gene. This is a 25,000-MW protein responsible for the surface exclusion (prevention of conjugative mating with related donors) and present as 20,000 copies in the outer membrane of *E. coli* carrying R6-5. M. Binns (St Louis) has demonstrated that the *traT* product of another conjugative

plasmid, R100, impairs the formation of the terminal complement complex. Furthermore, *traT*-containing cells remain resistant to complement even if they have been pretreated with antibodies to *traT*. This suggests that the *traT* protein does not actively protect bacteria against serum but has a passive effect, possibly involving subtle alteration of the architecture of the cell envelope. In addition, studies by Timmis of mutants in the *traT* gene indicate that the serum-resistance and surface-exclusion properties of *traT* protein may be distinct. Whatever the mechanisms involved, serum-resistance determinants, such as *traT*, may provide a selective advantage leading to the maintenance of conjugative resistance plasmids in pathogenic bacteria.

An important prerequisite for invasive pathogens is the ability to proliferate in the low free-iron conditions prevailing in animal tissues. A significant proportion of *E. coli* strains causing bacteraemia in humans or animals are found to carry ColV plasmids. In addition to encoding colicin V production and serum resistance, such plasmids specify accessory iron-sequestering mechanisms. P. Williams (Leicester) has shown that they determine an extracellular iron-chelating compound of low molecular weight which competes successfully with transferrin for iron *in vivo*. In addition, they specify, together with chromosomally determined components, an active iron uptake system. The possession of these extra mechanisms for acquiring free iron is thus sufficient to allow growth of *E. coli* in the tissues of infected animals.

Knowledge gained from *E. coli* should facilitate the analysis of more complex, chromosomally determined pathogenic mechanisms in less tractable organisms. And cloning of virulence determinants provides the promise of effective vaccines against enteric and other infections. □

J.R. Saunders is in the Department of Microbiology, University of Liverpool.



100 years ago

PHYSICAL NOTES

M. Plantamour continues to study with his sensitive levels the phenomena of periodic rise and fall of the ground which he has observed in Switzerland. He believes he has established a connection between these periods and those of the changes of temperature of the earth's surface, there being an annual change of level in an east-west direction corresponding with the mean temperatures of the surface during the year.

From *Nature* **23**, 31 March, 517, 1881.

*On the 'Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids' (organized by S.B. Levy, R.C. Clowes and E. Koenig), held from January 5th to 9th, 1981 in Santo Domingo, Dominican Republic.

Superluminal radio sources

from Guy Pooley

THERE was some understandable disbelief when it was reported about 10 years ago that some radio sources appeared to expand at velocities greater than that of light. Astronomers, though, are used to the impossible and since then there have been advances in both theory and observations. The article by Pearson *et al.* of the Cal Tech radio astronomy group (see this issue of *Nature*, p.365) describes observations of one particular quasar, 3C273, made with the US very long baseline interferometry (VLBI) network. They have applied the techniques of map-making developed by Readhead and Wilkinson at Cal Tech using so-called 'closure phase' measurements.

These observations all require the highest possible angular resolution — 1 milli-arcsecond or better — that at present can only be achieved by radio astronomers using transcontinental or inter-continental baselines. With these systems, it is not yet possible to measure absolute phase differences between signals arriving at the different stations; uncertainties in the clocks, geometry of the Earth, atmospheric paths and so on are too large. But if more than two receiving stations are available, there is some redundancy in the phase information, and the more stations in use the greater are the constraints that this places on the data. Pearson *et al.* used four, sometimes five stations and therefore six or ten baselines. Earlier long-baseline observers had to be content with fringe amplitude information only, and this was used to constrain models for the distribution of the radio sources on the sky; but it could not, for example, be used to distinguish between models which differed by a rotation of 180°.

The high apparent transverse velocities have been detected in several radio sources: quasars like 3C273, 3C279 and 3C345, and a galaxy, 3C120. These sources have been studied by VLBI techniques, until now without the use of closure phase. The results were perhaps most convincing in the case of 3C345, which could be modelled well by a simple double source and showed $v_{\text{apparent}} \approx 7c$.

3C273 is a source of some complexity. On the large scale, it has coincident radio and optical 'jets' some 20 seconds of arc (80 kiloparsecs) long. This jet is of course not detected in the long-baseline study, which shows the expanding feature some 0.01 seconds from the (presumed) central object. This feature is apparently moving with a transverse velocity of about 10c, rather more than had been estimated from the phaseless observations. The orientations of the large- and small-scale structures differ by about 20°.

One thing that cannot be done with closure phase is to find the absolute position of the source. It is therefore not yet possible to say with certainty which

features in the maps are 'stationary'.

The theoretical problems are not entirely new; Nova Persei (1901) gave similar difficulties, which were explained in terms of simultaneous excitation of separated clouds by a shock-front. Blandford, McKee and Rees reviewed the possibilities for the expansion of extragalactic radio sources (*Nature* 267, 211; 1977). The range of possibilities is large. There are the 'acausal' models such as the 'Christmas tree' version involving independent sources whose unrelated eruptions mimic expansion and there are models depending on the variation of propagation conditions either close to the source or in the intervening space (gravitational lenses, or scintillation in some intervening medium of galaxy). Neither of these general types seems to explain naturally the common features, particularly the fact that the sources always seem to expand.

The most popular models involve ejection of matter in some form of relativistic jet; there are several variants in which this can lead to high apparent transverse velocities. One particular attraction lies in the fact that the brightness temperatures of some compact radio sources, as implied by their rapid variations and inferred small sizes, are already uncomfortably high for the incoherent synchrotron process; a relativistic expansion eases this problem.

Scheuer and Readhead (*Nature* 277, 182; 1979) have developed one attractive version of this model. In it, they postulate that radio emitting objects are emitted repeatedly along one axis from the central object. Most of the light comes from the central, stationary object. The relativistic γ for the approaching component must range up to around 10, and any receding components would be practically undetectable. In their model, it is those quasars or galaxies whose axes are aligned almost along the line of sight for which dramatic effects are observed; those which are not so aligned are radio quiet, or at least radio dim, and may form central components of 'classical' double sources.

Pearson *et al.* make the point that the Scheuer-Readhead model is in some danger because of the dearth of radio-quiet quasars of comparable magnitude to that of 3C273; this is one of the questions for which careful statistics need assembling, and this can be done only by systematic mapping of expanding sources of this type. Further questions can now be asked; is the expansion uniform? When the size doubles, it might be surprising if the speed does remain constant. What are the statistics of a complete sample of quasars? Are the detailed structures consistent with the expanding jet models? The evidence is that rapid expansion is a very common feature among compact radio sources. □

Guy Pooley is in the Mullard Radio Astronomy Observatory, University of Cambridge.

DNA methylation and control of gene expression

from Tomas Lindahl

ABOUT 3% of cytosine residues in mammalian DNA are present as 5-methylcytosine, methylation of cytosine usually occurring soon after DNA synthesis. The existence of such signal groups has often stimulated speculation, but hard facts on their role have been scarce. However, recent experiments suggest a role for DNA methylation in regulating or blocking transcription of DNA. Such data further indicate that specific demethylation events may be necessary for gene expression during development and differentiation. A meeting on the subject was recently held at the University of Cologne* and one point that emerged was that an understanding of the role of DNA methylation in mammalian cells is unlikely to be obtained from studies on lower organisms and other model systems.

The relationship between DNA methylation and ribosomal gene function in several different organisms was reviewed

by A. Bird (MRC Mammalian Genome Unit, Edinburgh). Mouse cells carry about 100 ribosomal RNA genes per cell, occurring in two forms, unmethylated and heavily methylated. In various inbred mouse strains, between 5 and 30% of the sequences are methylated. Only the unmethylated ribosomal genes in mouse liver appear to be expressed, and the data point to an inverse correlation between DNA methylation and transcription. However, in *Xenopus laevis* all the chromosomal ribosomal DNA is heavily methylated, and this vertebrate clearly transcribes methylated sequences, although transcription may depend on the presence of an undermethylated region within the non-transcribed spacer between the ribosomal DNA genes. In contrast, *Drosophila* cells contain virtually undetectable amounts of 5-methylcytosine, so this organism is not likely to utilize DNA methylation for control of gene activity.

Many plant cells contain exceptionally high concentrations (up to 25%) of 5-methylcytosine in their DNA, and transcription of such methylated DNA does take

*A meeting on 'DNA Methylation and Genome Organization' (4-7 March, 1981) was organized by the Institute of Genetics, University of Cologne.

place (B. Deumling, Heidelberg). The reason for the heavy DNA methylation in plants remains unknown, but replacement of cytosine with a more radiation-resistant derivative could provide the cells with an increased tolerance to sunlight.

In animal cells, methylated cytosines are often present in relatively large amounts in non-transcribed satellite sequences, and fluctuation of the number of such regions in different cell types has often obscured more interesting variations in DNA methylation. This problem has been circumvented by examining the methylation of specific DNA sites within individual genes using restriction enzyme analysis and cloned DNA fragments as probes. Some of the most interesting recent data have come from work with DNA tumour viruses. L. Vardimon and W. Doerfler (Institute of Genetics, University of Cologne) showed for three different adenovirus type 2 transformed hamster cell lines, which contained the entire gene for the virus-coded 72K DNA binding protein, that the mRNA for the protein (as well as the protein itself) was present in only one of the cell lines. In the latter, the integrated viral DNA was unmethylated, as judged by its sensitivity to the *HpaII* restriction endonuclease, while the corresponding DNA sites in the two other cell lines were methylated. Similar results were obtained with a cloned viral DNA fragment methylated *in vitro* at a few cytosine residues. Furthermore, *in vitro* transcription studies (performed together with J. Corden, Strasbourg) indicated that the initiation of transcription of the gene for the 72K protein was severely inhibited by methylation of the DNA.

An oncogenic herpesvirus, Herpes saimiri, is similar to the adenoviruses in that the DNA in virus particles is completely unmethylated, while the largely non-expressed viral DNA present in virus-induced tumours or in transformed cell lines is heavily methylated (B. Fleckenstein and C. Kaschka-Dierich, Erlangen and R.C. Desrosiers, New England and Primate Research Center). As with the similar human lymphotropic herpesvirus, Epstein-Barr virus, the viral DNA in H.saimiri transformed cells is largely carried as multiple copies of nonintegrated plasmid DNA; S-adenosylhomocysteine treatment (which inhibits DNA methylation) markedly decreased the amount of viral DNA carried per cell.

As might be expected, there are exceptions to the rules. SV40 DNA from virions is apparently completely unmethylated. Microinjection experiments with SV40 DNA methylated *in vitro* with a highly purified rat liver methylase (which acted on all CpG sequences as well as several other C residues so that 2.5% of the cytosines were converted to 5-methylcytosine) seemed at variance with some of the data obtained for other tumour viruses (D. Simon and A. Graessmann, Berlin). On microinjection of 2 to 4 viral DNA molecules per host cell, there was apparently no

active demethylation of the SV40 DNA, but nevertheless, the methylated DNA induced SV40 determined antigens as effectively as nonmethylated DNA. Similar data were obtained with both permissive and nonpermissive host cells, so these experiments appear to suggest a difference in the sensitivity to DNA methylation between SV40 and larger DNA tumour viruses. The problems with this approach are that the incoming methylated DNA may have been replicated into nonmethylated daughter molecules, or the methylated DNA preparation may have contained a minority of unmethylated molecules which were transcribed.

An elegant system for the investigation of the effect of DNA methylation is provided by studying transformation of thymidine kinase deficient cells using cloned thymidine kinase genes. M. Wigler (Cold Spring Harbor Laboratory) used genes from Herpes simplex type I and from chicken, and compared the biological activity of unmethylated DNA and DNA methylated *in vitro* with *HpaII* methylase. The methylated genes showed distinctly lower transformation efficiency than the unmethylated controls, although the two kinds of molecules had the same frequency of incorporation. Methylation either in the promoter region or in the structural gene caused equal inhibition compared to unmethylated controls, so specific methylation of the promoter region did not seem to be of unique importance.

In the above experiments, the methylated cytosine residues of the incoming DNA, which were present in CpG sequences, appeared to serve as their own genetic determinants and were largely preserved for many generations. It seems clear that mammalian cells contain a maintenance methylase to retain the general methylation pattern of DNA. Similar data were presented by H. Cedar (Hebrew University, Jerusalem), who replaced dCTP with 5-methyl-dCTP in *in vitro* replication to obtain a very highly methylated DNA molecule. So far, only model experiments had been performed with the replicative form of Φ X174 DNA, containing one strand with all cytosines replaced by 5-methylcytosine. After introduction into hamster cells, where the DNA integrated at low efficiency, methylated sites at all CpG sequences seemed to be preserved while other 5-methylcytosines were replaced by unmodified cytosine on replication. This indicates that the maintenance methylase requires the CpG sequences for activity. However, the slow progress made in the characterization of mammalian DNA methylase(s) makes such observations seem somewhat preliminary.

Detailed restriction enzyme analysis of methylation sites within the human globin gene clusters (R.A. Flavell, National Institute for Medical Research, London) shows that many methylated residues persist also in genes that are being expressed, but globin DNA that is transcribed is

undermethylated compared to that present in sperm. These observations again seem in accord with the general idea that DNA methylation may play a role in controlling and preventing transcription. One way of overcoming inherent problems that arise from using DNA cloned in *Escherichia coli* in biological assays such as *in vitro* transcription is to use DNA heavily methylated *in vitro*, on the assumption that it is easier for a cell to specifically demethylate DNA than to methylate it, in analogy with the demethylation that apparently occurs with germ-line globin DNA on activation. It is not known how this demethylation is achieved, although several models are possible. For example, DNA may be replicated in a cellular compartment in which methylation at some sites does not occur, or specific DNA binding proteins may prevent the methylation at certain loci during replication. Alternatively, some 5-methylcytosine could be actively replaced by unmodified cytosine in the same way as misincorporated uracil residues are efficiently removed from DNA.

The role of DNA methylation in simpler organisms than mammalian cells is also being studied. In an important advance, R. Sager (Sidney Farber Cancer Institute, Boston) has shown that the maternal inheritance of chloroplast genes in *Chlamydomonas* results from the methylation of chloroplast DNA in female (but not in male) gametes, leading to the preferential degradation of the chloroplast DNA of male origin in zygotes. In cells carrying a *mat-1* mutation, the male chloroplast DNA also becomes methylated, and consequently the inheritance pattern is converted from maternal to biparental. M. Meselson (Harvard University) discussed the role of DNA adenine methylation for the direction of mismatch repair of phage λ heteroduplex molecules. Some, but not all, of various constructed mismatches having one strand methylated *in vitro* with a DNA methylase and the other strand unmethylated were susceptible to repair. While DNA sequence data on the regions around the mismatched residues are not yet available, it seems clear that the mismatches that are being repaired have to be present in a heteroduplex structure in which only one strand of the DNA is methylated, and that repair then occurs selectively in the unmethylated strand. Thus, neither fully methylated DNA nor unmethylated DNA is susceptible to mismatch repair, and it presumably only takes place in newly replicated DNA with the parental strand methylated. While mammalian cells probably also have the ability to perform mismatch repair, they do not seem to contain 6-methyladenine in their DNA, so it is an open question whether DNA methylation can guide mismatch repair events in higher organisms. □

Tomas Lindahl is Professor of Medical Biochemistry at the University of Gothenburg, Sweden, presently at the Imperial Cancer Research Fund Laboratories, London.

ARTICLES

Superluminal expansion of quasar 3C273

T. J. Pearson, S. C. Unwin, M. H. Cohen, R. P. Linfield, A. C. S. Readhead, G. A. Seielstad, R. S. Simon & R. C. Walker*

Owens Valley Radio Observatory, California Institute of Technology, Pasadena, California 91125, USA

Maps of the radio structure of 3C273 show directly that it expanded with an apparent velocity 10 times the speed of light from mid-1977 to at least mid-1980.

THE radioastronomical technique of very long baseline interferometry (VLBI) achieves an angular resolution better than 1 mas, equivalent to a physical resolution of the order of light-years in even the most distant galaxies and quasars. At this resolution, changes in the structures of such objects can be easily detected. In four extreme cases, three quasars (3C273, 3C279 and 3C345) and one galaxy (3C120) have been seen to expand with an apparent speed more than twice the speed of light, c (ref. 1). The evidence for such 'superluminal' expansion has been based on rather limited data, and the interpretation has been uncertain; but the recent development of the VLBI technique of hybrid mapping² allows us to map the radio brightness distribution to show the expansion directly and unambiguously.

Table 1 Observations

Epoch	Antennas*	Frequency† (MHz)	Flux density‡ (Jy)
1977.56	KGFO	10,651	40.9
1978.24	KGFO	10,651	44.5
1978.92	KGFO	10,651	34.9
1979.44	BKGFO	10,651	30.7
1980.27	BKG O	10,651	27.0
1980.52	BKGFO	10,651	27.8
1977.92	GFOH	5,011	46.2
1978.51	GFOH	5,011	45.6
1979.25	B G OH	4,997	40.0
1979.92	BKGFO	5,009	36.0

* B, 100-m, Max-Planck-Institut für Radioastronomie, Effelsberg, FRG; K, 36-m, Haystack Observatory, Massachusetts; G, 43-m, National Radio Astronomy Observatory, Green Bank, West Virginia; F, 26-m, Harvard Radio Astronomy Station, Fort Davis, Texas; O, 40-m, Owens Valley Radio Observatory, Big Pine, California; H, 26-m, Hat Creek Radio Astronomy Observatory, Cassel, California.

† The centre frequency of the 2-MHz band.

‡ Measured during the observation; typical uncertainty 0.5 Jy.

The 10.65-GHz observations were made in left-circularly polarized radiation, and the 5.0-GHz observations were made in linearly polarized radiation with the E-vector in PA 90°.

This article presents maps of the quasar 3C273 which show that it was expanding with an apparent velocity $v = (9.6 \pm 0.8)c$ between 1977.5 and 1980.5. (We assume $H_0 = 55 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.05$ throughout.)

3C273 is a 13-mag quasar with a redshift of 0.158 (ref. 3). An optical jet extends from a distance of ~ 10 arcs to ~ 25 arcs from the quasar in position angle (PA) -137° (ref. 4). The projected length of the jet is 84 kpc. The radio emission comes principally from two locations: a compact component (3C273B) coincident with the quasar, and a more diffuse component

(3C273A) near the end of the jet⁵⁻⁷. The compact component 3C273B consists of a flat-spectrum core with a steeper-spectrum extension in PA -117° (ref. 8). This direction is approximately that of the superluminal expansion. As the angular scale is increased, the predominant position angle rotates until it is more closely aligned with the optical jet⁹. Recent observations of 3C273 in all wavebands have been reviewed by Ulrich¹⁰.

3C273B was one of the first sources to be studied with VLBI^{11,12} and has been the subject of many subsequent observations. Changes in the visibility function were first observed in 1971^{13,14}. Although these observations were confined to a single baseline, it was immediately clear that the changes could be explained by a simple expansion. The data permitted other explanations, however, including models in which the centres of emission remained fixed while varying in intensity¹⁵. With more complete visibility data¹⁶⁻¹⁸, it became clear that the structure was more complex than a simple linear model of two or three components, and that the overall extent of the brightness distribution continued to increase. To avoid some of the ambiguities inherent in model-fitting, Cohen *et al.*¹ examined the visibility function measured between 1970 and 1977, and found that the first minimum (that at shortest projected baseline) in the visibility amplitude moved steadily towards smaller baselines, indicating a steady increase in the size of the source. They took as a measure of the angular size the quantity θ^* , defined as $1/(2w)$, where w is the projected baseline (in wavelengths) at the first minimum. If the source were a simple double, θ^* would be the separation of the two components. A later analysis¹⁹ of all the data available up to 1977.41 showed that the time-dependence of θ^* approximated a straight line with a slope of $0.41 \pm 0.04 \text{ m arc s yr}^{-1}$, corresponding to an apparent linear expansion rate of $v/c = 5.2 \pm 0.5$. This steady increase in θ^* (Fig. 16 of ref. 19) has provided the strongest evidence for superluminal expansion in 3C273. We know that 3C273 was not accurately represented by a simple double model during the period spanned by these observations, so θ^* was not a true measure of its angular size. We should not be surprised therefore that the maps presented here show that the true expansion speed differs from that inferred from the measurements of θ^* .

Maps

Since 1977 we have been monitoring 3C273 at two frequencies, 10.65 GHz and 5.0 GHz, using four or five antennas. The new observations are better than the old for several reasons: (1) we have used more antennas, sampling the visibility function at more spatial frequencies; (2) we have when possible included an antenna in Europe in addition to American antennas, thereby improving the resolution; (3) we have measured closure phase²⁰ in addition to visibility amplitude. These improvements enable us to make hybrid maps of the brightness distribution². The increase in the quantity of data has eliminated the ambiguities of model-fitting. The first maps at these two frequencies have already been published^{8,21}, and we have now observed a total of six epochs at 10.65 GHz and four epochs at 5.0 GHz (Table 1).

* Present address: National Radio Astronomy Observatory, Edgemont Road, Charlottesville, Virginia 22901, USA.

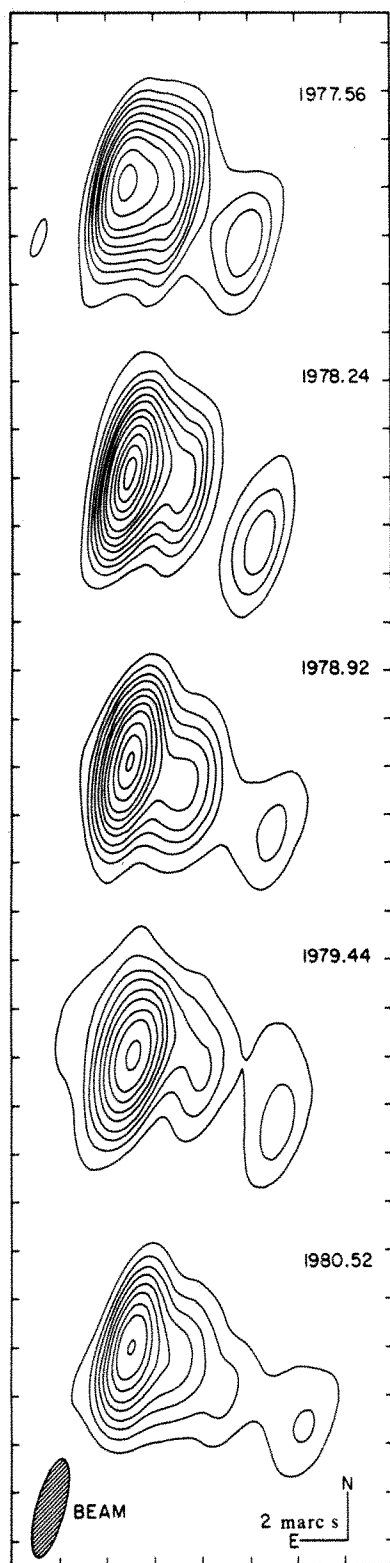


Fig. 1 Maps of 3C273B at 10.65 GHz at five epochs (see Table 1). The same restoring beam (an elliptical gaussian with FWHM 4.2×1.2 marc s, with the major axis in PA -15° , indicated by the hatched ellipse) is used for all the maps. The dynamic range of the maps is about 20:1. The contour levels for all the maps are 1.2, 2.4, 3.6, 5, 7, 9, 12, 15, 20, 25, 30 and 35×10^9 K. The absolute position of the source cannot be accurately determined from the data; here the maps have been lined up on the sharp eastern edge to show the motion of the weaker western component. The map for epoch 1977.56 is from ref. 8, which gives more details of the map-making procedure.

Full details of these observations will be published elsewhere; here we describe only the 10.65-GHz maps and the superluminal expansion that they reveal.

Figure 1 shows maps made from five of the six 10.65-GHz observations. It was not possible to make a reliable map from the observation of 1980.27 owing to a failure at one of the antennas. The beam (hatched ellipse in Fig. 1) is very elongated because 3C273 is close to the celestial equator, where the presently available network of antennas has very poor north-south resolution. Fortunately the principal extension of the source is close to the direction of maximum resolution. Two of the experiments included an antenna in Europe as well as those in America (see Table 1). For these, the east-west resolution was improved by a factor of two. To make the maps directly comparable, however, the restoring beam appropriate to the lowest-resolution observations was used for all the maps.

The five maps in Fig. 1 span a period of 3 yr. All the maps show a bright eastern peak and a lower-brightness extension to the west. In all cases there is a local maximum in the western

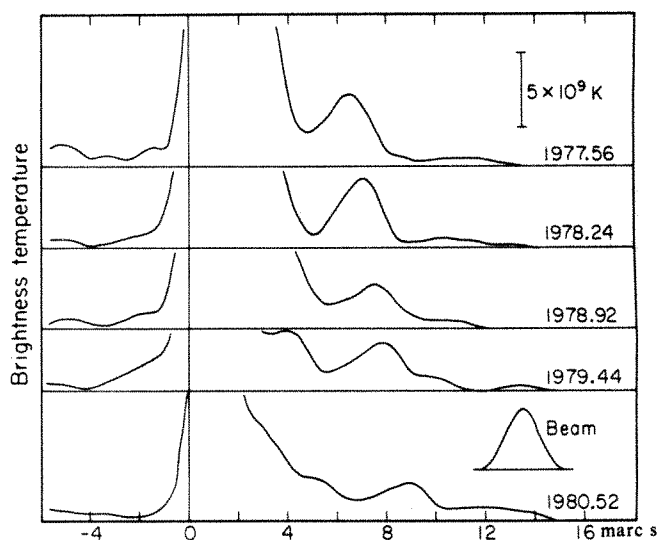


Fig. 2 Brightness-temperature profiles through the maps of Fig. 1 along the principal axis of the source (PA -116°). The profiles have been truncated to show the evolution of the weak western component, and they have been aligned at the half-power point on the eastern edge, labelled 0 on the abscissa. The vertical displacement between the profiles is proportional to the difference in epoch. The profile of the restoring beam (FWHM 1.42 marc s) is indicated.

extension between 6 and 8 marc s from the main peak. The most striking difference between the maps is the steady increase in the separation of this local maximum from the main peak. The position angle defined by the main peak and the secondary maximum has not changed significantly; in all the maps it is $-116^\circ \pm 2^\circ$. The maps also show that the source is not straight but curves to the south with increasing separation from the main peak. This phenomenon has been noted previously^{8,9,18}.

Expansion velocity

The profiles in Fig. 2 show that the separation of the western maximum from the main peak increased from ~ 6 marc s to ~ 8 marc s between 1977.5 and 1980.5. The time-dependence of the separation, θ , is shown in Fig. 3. In addition to the 10.65-GHz points, one 5-GHz measurement is included to show that the size of the source is the same at both frequencies. (The other 5-GHz observations do not have sufficient resolution to be directly comparable, but they are consistent with this conclusion.) θ was determined by measuring the separation of the peak of the western maximum from the half-power point in the sharp eastern edge, and then subtracting half the FWHM of the restoring beam (0.71 marc s) to correct for resolution. This measure is less affected by changes in the internal structure of the eastern component than a separation measured to the peak

of the eastern component. In all cases the measurement was made along a line with PA -116° .

The values of θ all fall remarkably close to a straight line in Fig. 3, indicating that the expansion is linear, with no evidence for acceleration or deceleration. The angular expansion rate, derived by a least-squares fit of a straight line (omitting the 5.0-GHz point), is 0.76 ± 0.04 marc s yr^{-1} , corresponding to a linear expansion rate of $v/c = 9.6 \pm 0.5$. The extrapolated intercept $\theta = 0$ is at 1970.1 ± 0.6 . If the expansion speed has not changed, one might expect to see an increase in total flux density at this epoch, corresponding to the formation of the new component. No such increase was seen²². The flux density of 3C273 has irregular variations with a time scale of one or two years, but it is not possible to correlate changes in the flux density with structural changes.

θ is larger than the values of θ^* extrapolated from the older data, but this is due to the difference in measuring technique. θ^* , being determined directly from the visibility function, is not related in a simple way to the brightness distribution except when the source is a simple double. If the expansion were conformal (if the brightness distribution retained the same shape but simply increased in scale), θ would differ from θ^* by a constant factor, but the real behaviour is more complicated.

Acceleration?

The angular size of 3C273B derived by extrapolating the expansion rate shown in Fig. 3 back to the early 1970s is smaller than the size then observed. This implies that either the expansion speed has increased, or the expansion has not been conformal. It is possible to devise non-conformal models that can explain the qualitative features of the observations without requiring an acceleration. In triple models, unlike doubles, the positions of the minima depend not only on the separations of the components but also on their relative intensities. One such model involves one fixed component and two components both moving with the same velocity. Suppose that at the earlier epochs (1970–75) the source consisted of three principal compact components: a core C at the eastern end, with flux density S_C , and two components A and B moving away from the core at $10c$ in PA -116° , with flux densities S_A and S_B . Let A be the more distant component. When the separation of A from C is much greater than the separation of B from C, the first visibility minimum measures the distance between A and the centroid of B and C. The expansion rate derived from θ^* is then $10cS_C/(S_B + S_C)$. This rate was $5.6c$ (ref. 19), implying that $S_B \approx S_C$ in 1970–75. Component A has since decayed below the level of detectability, and the new maps measure the separation of B from C which is increasing at $10c$. Of course, this model is an oversimplification, as the maps show additional components lying between B and C.

It is not possible to eliminate models involving acceleration, but a constant-velocity model seems to be favoured by the constancy of the expansion rate shown by the maps of Fig. 1. If the constant-velocity model is correct, we would expect further components to appear in maps made during the next few years.

Relativistic jet

Although several explanations have been suggested for superluminal expansion^{23,24} we shall discuss only the model²⁵ in which the expansion is presumed to represent a physical motion almost directly towards the observer; we are looking down a jet of relativistic material expelled from the quasar. The eastern end of 3C273B is the 'core', the point at which the jet becomes optically thick. In a steady flow, this point remains fixed even though the radiating material has a high velocity. The moving (western) component is a 'knot' propagating outwards along the jet. At large distances, the jet becomes visible as the optical jet, and the region of impact on the external medium is represented by the radio source 3C273A. This model accounts for (1) the milli-arcsecond morphology of 3C273B; (2) the fact that the expansion is roughly along the source axis; and (3) the approximate alignment between the axis of 3C273B and the optical jet.

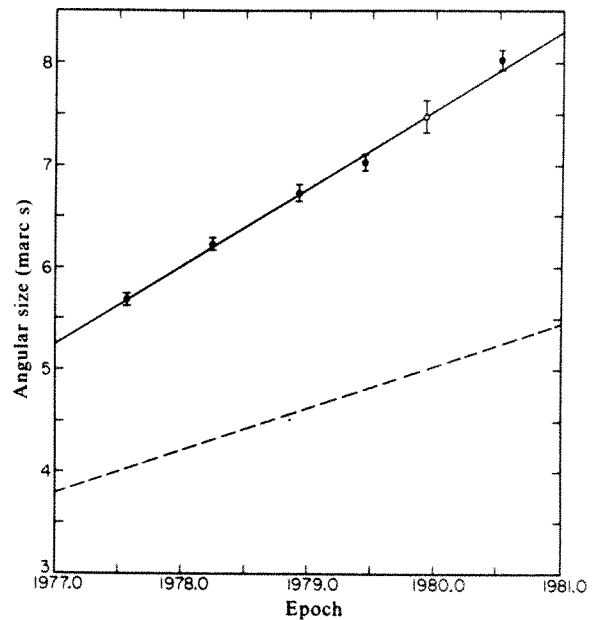


Fig. 3 Angular size of 3C273 as a function of time, measured as described in the text. Five 10.65-GHz measurements (●) and one 5.0-GHz measurement from 1979.92 (○) are shown. The dashed line at the bottom is an extrapolation of the expansion rate derived from observations of θ^* before 1977 (ref. 19). The solid line is a least-squares fit to the 10.65-GHz measurements, and has a slope of 0.76 marc s yr^{-1} .

(The difference of 20° between the radio axis and the optical axis is attributed to curvature in the jet⁹.) If the constant-velocity interpretation of the older data is correct, then at least two knots have been ejected from the core in a period of ~ 10 yr. Alternatively, an accelerating expansion can be accommodated in this model: it could be due to real acceleration in the flow along the jet or to a curvature of the jet.

To account for an apparent transverse velocity v as large as $9.6c$, the angle ϕ between the jet and the line of sight must be small. The apparent velocity v is related to ϕ and to v_0 , the space velocity of the moving feature, by

$$v = v_0 \sin \phi / \{1 - (v_0/c) \cos \phi\}$$

(ref. 26). If $v = 9.6c$, this equation requires that $\phi < 11.9^\circ$ ($\cos \phi > (v^2 - c^2)/(v^2 + c^2)$) if $v_0 < c$, and v_0 is minimized when $\phi = 6.0^\circ$. The probability that ϕ should be as small as 11.9° in a randomly selected quasar is 1%, or 2% if the jet is two-sided.

3C273 is far from being a randomly chosen quasar: it is the closest quasar in the 3CR sample²⁷ and has an intrinsic optical intensity ~ 4 times the intensity of the other quasars in the sample. It is surprising that it should also be exceptional in having a jet pointing almost directly at Earth. A possible explanation is that the optical radiation, like the radio, is not isotropic. There is no direct evidence for this, but beaming of the optical continuum cannot be ruled out. If the optical radiation is isotropic, then the small angle between the radio jet and the line of sight causes grave difficulties for the simple relativistic theories. For example, Scheuer and Readhead²⁸ attribute the difference in radio intensity between optically selected and radio-selected quasars to the relativistic beaming of the radio radiation—radio-selected quasars point towards Earth. We would expect on this model to see 50–100 radio-quiet quasars of comparable optical magnitude to 3C273, which we do not. It may be necessary to consider other models in which the radio radiation is not confined to so small an angle; one such model has been suggested by Rees²⁹.

Conclusion

The maps presented in Fig. 1 provide the first direct and unambiguous evidence of superluminal expansion in any radio source. For the period of 3 yr spanned by the observations the superluminal expansion could be attributed to the movement of

a single knot away from the nucleus along the jet. The apparent constant velocity of 9.6c observed in this period is an important constraint on theories of apparent superluminal expansion. Further observations of both 3C273 and other superluminal sources are needed to show whether the expansion is always at a constant velocity; the present results show that such analyses must be based on maps, rather than on the motion of features in the visibility function.

The observations described here were made under the auspices of the United States VLBI Network Users Group. We

Received 3 February; accepted 25 February 1981.

1. Cohen, M. H. *et al.* *Nature* **268**, 405–409 (1977).
2. Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* **223**, 25–36 (1978).
3. Schmidt, M. *Nature* **197**, 1040 (1963).
4. Greenstein, J. L. & Schmidt, M. *Astrophys. J.* **140**, 1–34 (1964).
5. Hazard, C., Mackey, M. B. & Shimmins, A. J. *Nature* **197**, 1037–1039 (1963).
6. Conway, R. G. & Stannard, D. *Nature* **255**, 310–312 (1975).
7. Perley, R. A., Fomalont, E. B. & Johnston, K. J. *Astr. J.* **85**, 649–658 (1980).
8. Readhead, A. C. S. *et al.* *Astrophys. J.* **231**, 299–306 (1979).
9. Readhead, A. C. S. *et al.* *Nature* **276**, 768–771 (1978).
10. Ulrich, M.-H. *Space Sci. Rev.* (in the press).
11. Broten, N. W. *et al.* *Nature* **215**, 38 (1967).
12. Clark, B. G., Cohen, M. H. & Jauncey, D. L. *Astrophys. J. Lett.* **149**, L151–L152 (1967).
13. Whitney, A. R. *et al.* *Science* **173**, 225–229 (1971).
14. Cohen, M. H. *et al.* *Astrophys. J.* **170**, 207–217 (1971).

thank the many people involved in scheduling the observations, recording the data at the participating observatories, and processing the tapes, for indispensable contributions to this paper. The work was supported by the NSF via grants AST 77-00247 and AST 79-13249 to the Owens Valley Radio Observatory; the National Radio Astronomy Observatory is operated by Associated Universities, Inc., under contract with the NSF; the astronomical VLBI programs at Harvard Radio Astronomy Station, Hat Creek Radio Astronomy Observatory and Haystack Observatory are supported by the NSF.

15. Kellermann, K. I. *et al.* *Astrophys. J. Lett.* **189**, L19–L22 (1974).
16. Legg, T. H. *et al.* *Astrophys. J.* **211**, 21–30 (1977).
17. Schilizzi, R. T. *et al.* *Astrophys. J.* **201**, 263–274 (1975).
18. Niell, A. E., Kellermann, K. I., Clark, B. G. & Shaffer, D. B. *Astrophys. J. Lett.* **197**, L109–L112 (1975).
19. Seielstad, G. A. *et al.* *Astrophys. J.* **229**, 53–72 (1979).
20. Rogers, A. E. E. *et al.* *Astrophys. J.* **193**, 293–301 (1974).
21. Cohen, M. H. *et al.* *Astrophys. J.* **231**, 293–298 (1979).
22. Andrew, B. H., MacLeod, J. M., Harvey, G. A. & Medd, W. J. *Astr. J.* **83**, 863–899 (1978).
23. Blandford, R. D., McKee, C. F. & Rees, M. J. *Nature* **267**, 211–216 (1977).
24. Marscher, A. P. & Scott, J. S. *Publ. astr. Soc. Pacif.* **92**, 127–133 (1980).
25. Blandford, R. D. & Königl, A. *Astrophys. J.* **232**, 34–48 (1979).
26. Ginzburg, V. L. & Syrovatskii, S. I. *A. Rev. Astr. Astrophys.* **7**, 375–420 (1969).
27. Schmidt, M. *Astrophys. J.* **151**, 393–409 (1968).
28. Scheuer, P. A. G. & Readhead, A. C. S. *Nature* **277**, 182–185 (1979).
29. Rees, M. J. *Proc. IAU Symp.* No. 94 (in the press).

Human immunoglobulin κ light-chain genes are deleted or rearranged in λ -producing B cells

Philip A. Hieter*, Stanley J. Korsmeyer†, Thomas A. Waldmann† & Philip Leder*

* Laboratory of Molecular Genetics, National Institute of Child Health and Human Development,
† Metabolism Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

We have examined 10 human λ light-chain-producing B lymphocytes and found that genes for the κ constant region have been deleted in each. Only one of the 20 possible κ alleles in these cells has been retained and even this exception is rearranged. In contrast, λ constant region genes remain in the germ-line configuration in each of the eight κ light-chain-producing human B lymphocytes examined. Such observations suggest a hierarchy of light-chain-gene rearrangement beginning with κ and proceeding to λ .

A NUMBER of choices are involved in the differentiation of an antibody-producing lymphocyte. For example, usually only one light-chain gene is functionally expressed in a given cell, the choice having been made between the two light-chain classes (κ or λ) and their respective alleles¹. In addition, after commitment to the expression of a particular light-chain, the phenotypically excluded class and allele apparently remain silent throughout the functional life of the lymphocyte. These phenomena are called isotypic (κ versus λ) and allelic (one κ allele versus the other) exclusion. Whether they are tightly programmed or stochastic remains an extremely interesting but unanswered question.

Fortunately much of what we know about the formation of an active mouse κ gene is relevant to the question of allelic exclusion. In some B (antibody-producing) cells, the silent allele remains in the germ-line configuration and has not undergone the somatic V/J rearrangement necessary to form an active gene^{2–4}. In other cells, the silent allele has undergone V/J recombination but, because of the variability of this reaction, produces an out-of-phase J sequence and, consequently, a nonsense gene^{5,6}. In still other cells, recombination occurs at other than a J region making it impossible to form a coherent coding sequence^{7,8}. In each of these instances, the molecular basis for the silence of the excluded allele is clear, but the basis for the selection and control of this process, if any, remains obscure.

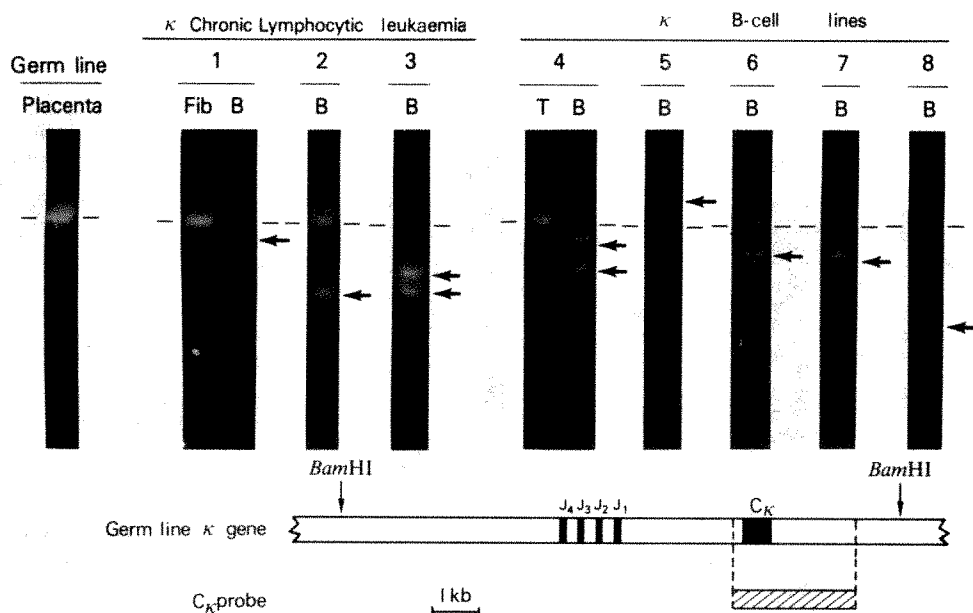
Little is known of the nature of isotypic exclusion, because studies of λ/κ expression in the mouse are necessarily limited by the fact that the λ class is expressed at low frequency in this organism. Recently, however, Alt *et al.*⁹ examined three long-

established λ chain-producing mouse myeloma lines and found that κ genes had been rearranged in two of them and possibly lost in a third. This observation led them to a proposal (see below) regarding the control of allelic and isotypic exclusion by the formation of functional genes. This proposal can be usefully pursued in man. In contrast to the mouse, human immunoglobulins are composed of one-third λ light-chain-bearing molecules and, therefore, the human λ gene system is likely to contribute significantly to antibody diversity. Accordingly, we have turned to man and an array of established B-cell lines and freshly obtained leukaemic cells to ask whether the structure and arrangement of these genes offer some clue as to the mechanism or order of their selection.

We show here that at least one light-chain gene in each light-chain-producing B cell has undergone rearrangement while its allelic partner has either remained in the germ-line configuration, undergone a second rearrangement or been lost. In each cell expressing a κ gene, the λ genes are present in their germ-line configuration, but in surprising contrast, in each λ -producing cell, κ genes have been either deleted or undergone a rearrangement. This consistent correlation suggests, among other possibilities, a putative hierarchy of light-chain-gene rearrangements in which the process is ordered so that it usually progresses from the κ to the λ genes.

κ genes are rearranged and deleted in κ -expressing cells

The arrangement of the κ light-chain genes in human κ -expressing B cells was determined using a radioactive DNA



autoradiograms were developed. The germ-line (embryonic) κ gene map demonstrating the 12-kilobase *Bam*HI C_{κ} -containing fragment and the 2.5-kilobase *Eco*RI embryonic C_{κ} -containing fragment (hatched bar) which was used as the C_{κ} probe are shown. The four human J_{κ} gene segments identified by heteroduplex analysis¹⁰ are indicated.

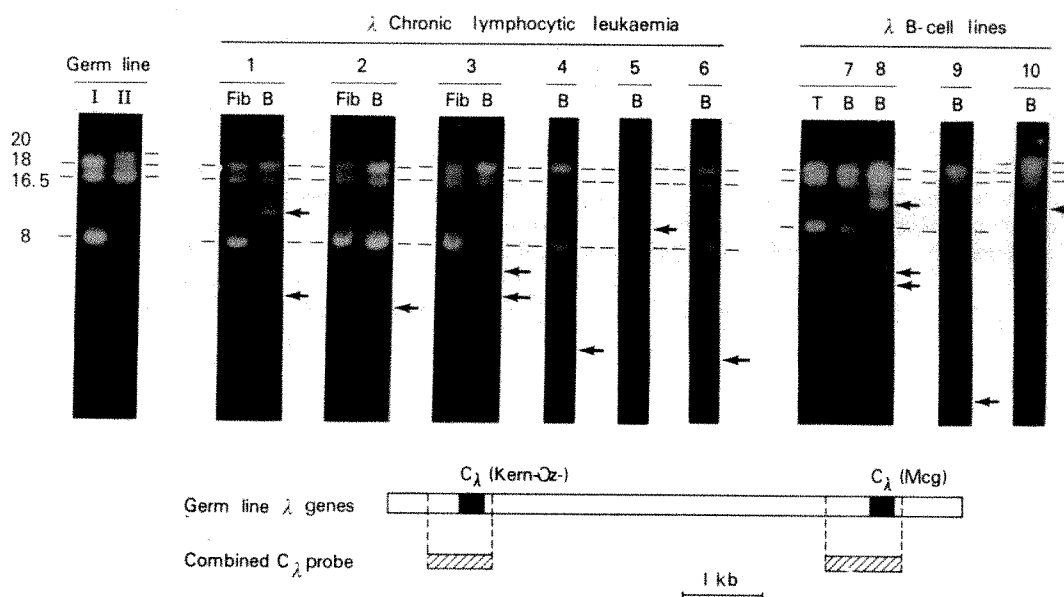
fragment containing the human κ constant region gene¹⁰ as a hybridization probe to detect this gene among fragments generated by *Bam*HI endonuclease digestion of genomic DNA (Fig. 1). Non-rearranged genomic DNA yields a single *Bam*HI fragment containing the single C_{κ} and four human J_{κ} segments¹⁰. Examination of such *Bam*HI fragments from non-B cells (placental tissue, fibroblasts or T cells) derived from 24 different individuals (see representative controls in Figs 1 and 4) showed them all to be in the germ-line configuration. Human κ light-chain-expressing B cells (including lymphocytes derived from three high-count chronic lymphocytic leukaemia patients, three B-cell lines established spontaneously from κ -expressing B-cell malignancies, and two κ -secreting Epstein-Barr virus-transformed human B-cell lines) were then examined. All eight κ -expressing B cells demonstrated at least one *Bam*HI fragment rearrangement when compared with the germ-line configura-

tion (Fig. 1). The second C_{κ} allele remained in the embryonic configuration in four cases, showed a DNA rearrangement in two cases and was deleted in two cases.

λ genes are rearranged in λ -expressing B cells

Analyses of human myeloma proteins¹¹ demonstrated that the λ constant region consists of at least four non-allelic forms in man which bear the non-allelic markers, Kern⁺Oz⁻, Kern⁻Oz⁺, Kern⁻Oz⁻ and Mcg. To identify these genes in λ chain-producing cells, we used a C_{λ} probe consisting of a mixture of two ³²P-labelled DNA fragments derived from cloned human λ constant genes (P.A.H., unpublished results). These were a 1.2-kilobase *Bam*HI-*Eco*RI fragment bearing the C_{λ} (Kern⁻Oz⁻) gene and a 0.8-kilobase *Bam*HI-*Hind*III fragment bearing the C_{λ} (Mcg) gene (Fig. 2, bottom). Because of the close amino acid sequence

Fig. 2 Detection of rearranged λ genes in λ -expressing human B cells. λ -Expressing B cells examined included the monoclonal λ light-chain-bearing B lymphocytes from patients (cases 1-6) with high-count chronic lymphocytic leukaemia, as well as four long-term λ -secreting Epstein-Barr virus-transformed normal human B-cell lines (cases 7, 8, 9, 10). Fibroblast cell cultures from cases 1, 2, 3 and spontaneous T-cell lines from cases 7 and 8 served as matched controls in the germ-line configuration. The two observed germ-line patterns of C_{λ} gene fragments are presented and referred to as type I and type II. As before, genomic DNA from all cells was digested with *Eco*RI, size fractionated by agarose gel electrophoresis, transferred to DBM paper and hybridized with ³²P-labelled combined C_{λ} probe. The C_{λ} probe (hatched bars) was a combination of the 1.2-kilobase embryonic *Bam*HI-*Eco*RI fragment containing the Kern⁻Oz⁻ C_{λ} gene and the 0.8-kilobase embryonic *Bam*HI-*Hind*III fragment containing the Mcg C_{λ} gene (as evidenced by nucleotide sequence analysis; P.A.H., unpublished results).



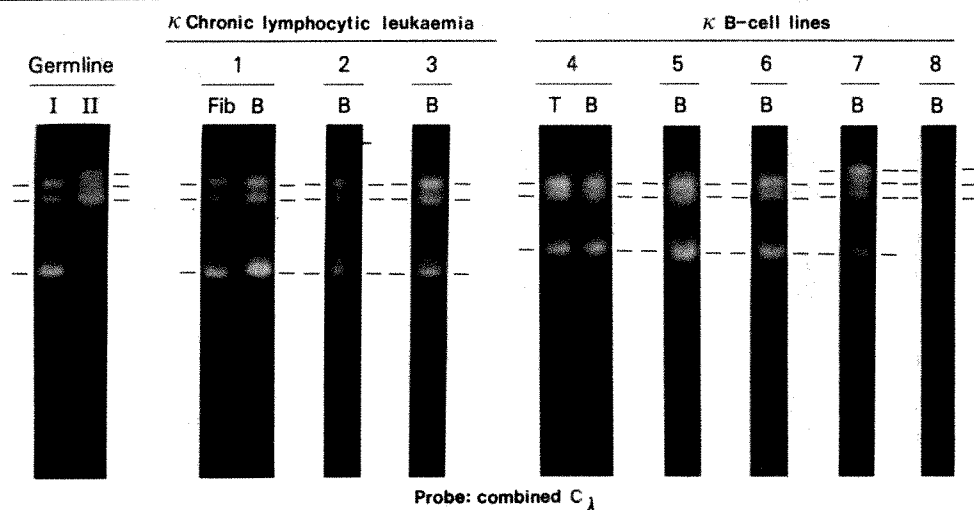


Fig. 3 Detection of λ genes in germ-line configuration in κ -expressing B cells. DNA obtained from the eight κ -expressing B cells shown in Fig. 1 and the available matched germ-line controls were digested with *Eco*RI, size fractionated by agarose gel electrophoresis, transferred to DBM paper and hybridized with the combined C_λ probe (described in Fig. 2). Cases 1–6 show type I germ-line λ genes; case 8, type II germ-line λ genes. We presume that case 7 is heterozygous for both the 8- and 20-kilobase *Eco*RI site alleles, although no matched germ-line control was available for this B-cell line.

homology of the known human λ constant regions, the combined C_λ probe would be expected to cross-hybridize to all four of the C_λ genes. Analyses of *Eco*RI digests of 24 non-B-cell genomic DNAs (placenta, fibroblast and T cells), which were carried out to establish the germ-line pattern, demonstrated a genetic polymorphism in *Eco*RI endonuclease sites that produced two predominant fragment patterns (Fig. 2, germ-line lanes), one which yields 8-, 16.5- and 18-kilobase fragments and another which yields 16.5-, 18- and 20-kilobase fragments. Such polymorphisms, which seem quite common at this locus (P.A.H., unpublished), provide useful chromosomal markers for further genetic studies.

Keeping in mind that apparent alterations in the *Eco*RI fragment pattern at the λ locus might be due to these polymorphisms, we examined a number of human λ chain-expressing B lymphocytes. These included monoclonal B cells from six patients with chronic lymphocytic leukaemia and four normal human B-cell lines transformed by Epstein-Barr virus. *Eco*RI digests of all 10 λ -expressing B cells showed at least one rearranged C_λ allele that was not seen in the matched non-B cells (either fibroblasts or T cells) from the same individual. Three of the B cells displayed double rearrangements that could be used to establish the order of the λ genes in chromosomal DNA. This was done by noting that in three cases (1, 3 and 8; Fig. 2), the 8-kilobase *Eco*RI fragment was lost in a double rearrangement of λ alleles, while the 16.5- and 18-kilobase fragments were preserved. Because V/J recombination occurs to the 5' side of an expressed constant region and involves deletion of DNA between V and J regions^{12,13,19}, the retained constant λ genes must be 3' to those on the 8-kilobase fragment.

λ genes are in the germ-line configuration in κ -expressing B cells

The arrangement of the λ light-chain genes in the eight human κ light-chain-expressing B cells analysed above is shown in Fig. 3. In all eight of the κ -expressing B cells, the *Eco*RI λ gene-

containing fragments remain in the germ-line configuration. Note that case 7, which demonstrates both the 8- and 20-kilobase *Eco*RI restriction fragments, seems to be heterozygous for the type I and type II polymorphic gene configurations.

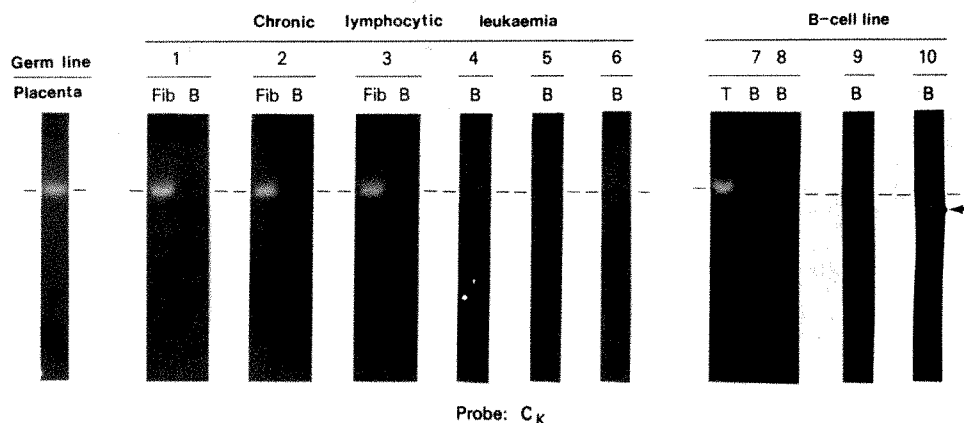
κ genes are deleted or rearranged in λ -expressing B cells

In clear contrast to the germ-line arrangement of λ genes in κ -expressing cells, both C_κ alleles are deleted in all six λ -expressing leukaemic B cells (*Bam*HI digest shown in Fig. 4). In addition, both C_κ alleles were deleted in three of the four λ -expressing B-cell lines. In the other case, one κ allele was rearranged while the other was deleted (case 10, Fig. 4). Although case 10 was a long-term cell line, it was not physically cloned. The finding of C_κ gene deletions in the nine λ -expressing B cells was confirmed in the *Eco*RI digests of genomic DNA from these same cells (data not shown). In contrast to the observed deletion in the clonal B cells, the C_κ gene was present in the germ-line configuration in *Bam*HI and *Eco*RI digests of DNA from matched (same individual) T cells or fibroblast lines, demonstrating that the κ gene deletions are apparently limited to the λ -expressing B cells (Fig. 4).

The κ gene deletions usually include the J segment

To examine the possibility that the C_κ deletion could extend to J and V region coding segments, conceivably involving the entire chromosome, we prepared probes corresponding to four J_κ segments and to a V-region gene (see Fig. 5) and tested for complementary sequences in each cell line. In the representative case shown (Fig. 5), C_κ and J_κ coding segments fail to hybridize to a complementary sequence within the λ -expressing B cell, indicating that the deletion has included the J_κ segment. However, the V-region probe detects complementary fragments indicating that this V region family has not been included in the deletion of C_κ and J_κ . The presence of V_κ sequences demon-

Fig. 4 C_κ gene deletion or rearrangement in λ -expressing B cells. The 10 λ -expressing B cells (B) and their available matched germ-line control fibroblasts (Fib) or T cells (T) were *Bam*HI-digested as before and then hybridized with the ³²P-labelled 2.5-kilobase *Eco*RI C_κ gene fragment (see Fig. 1). We see very faint bands present in several cases of chronic lymphocytic leukaemia B cells that are not evident in the figure. These represent the germ-line C_κ segments present in the 3–15% contaminating monocytes and T cells found in leukaemic B-cell preparations and are absent from all the cultured cell lines.



strates that the deletion event is localized and does not represent merely a loss of the chromosome bearing the κ locus. Identical studies of the remaining deletion examples gave identical results with a single exception—in one case, a J_{κ} was present, but in a rearranged configuration (data not shown). This exception nevertheless suggests that such deletions can occur from various points in the chromosomal DNA.

Several gene configurations can account for allelic and isotype exclusion

It is already clear from various analyses in the mouse that a significant proportion of κ light-chain-gene rearrangements are aberrant; that is, they create non-functional or invalid genes, genes incapable of producing normal light-chains⁵⁻⁸. Multiple rearrangements are not uncommon in mouse myeloma cells^{5-8,14-18}. Some of these rearrangements seem to depend on the normal V/J recombination mechanism that occasionally creates a nonsense sequence and, hence, a null gene^{5,6}. Other rearrangements do not seem to depend on V/J recombination, but involve additional mechanisms (conceivably those analogous to the heavy-chain switch) that produce aberrant recombinants between regions that may or may not encode authentic V or J segments^{7,8}. One such rearrangement actually involves a

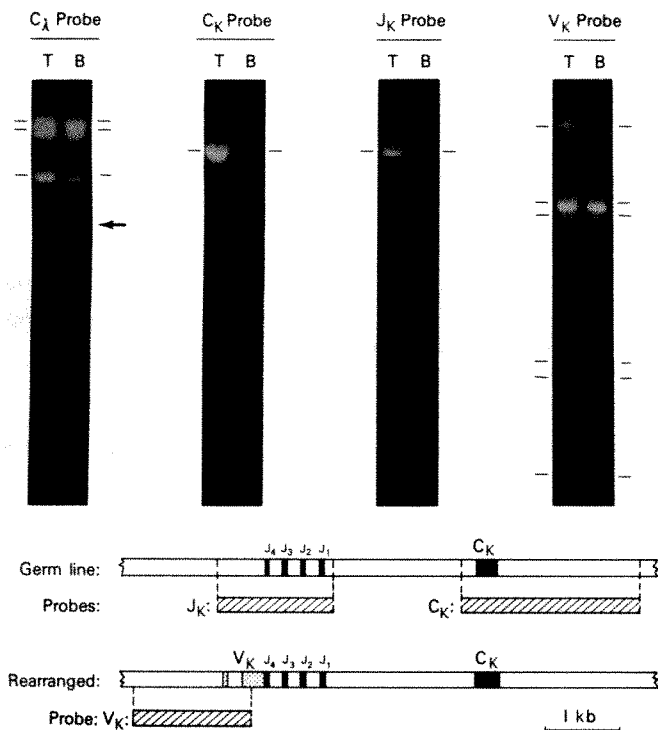


Fig. 5 Extent of κ gene deletion in λ -expressing B cells. A comparison of C_{λ} , C_{κ} , J_{κ} and V_{κ} gene segments in a representative λ -expressing B cell (case 7, Fig. 2) and its matching (same individual) T cell is shown. The probes are shown as hatched bars under each map. ^{32}P - C_{λ} and ^{32}P - V_{κ} (see below) probes (shown in figure) were hybridized to the *Eco*RI digests of B cells and available germ-line controls. ^{32}P - J_{κ} and ^{32}P - C_{κ} probes were hybridized with the *Bam*HI digests of the same cells. The V_{κ} gene probe was the 1.6-kilobase *Hpa*II fragment derived from the active V_{κ} - J_{κ} recombinant of the κ -expressing cell line RPMI 6410¹⁰. This fragment includes 63 bases of the 5' variable region corresponding to the first framework (sequence not shown). The amino acid translation of the nucleotide sequence yielded a protein sequence identical to that obtained for the human light-chain MAG²³ which is a member of κ subgroup II²⁴. The J_{κ} gene probe was a 1.7-kilobase germ-line (embryonic) *Cfo*I restriction fragment containing four human J_{κ} segments. The C_{κ} probe is described in Fig. 1 legend; the C_{λ} probe, in Fig. 2 legend.

recombinational 'hotspot' between the mouse J and C regions that is complementary to an extension of the small palindromic sequence thought to be involved in V/J recombination⁷. It is not known what proportion of DNA rearrangements during B-cell development result in the formation of an active gene, but it is conceivable that all light-chain alleles of both classes will occasionally rearrange in a single cell to form non-functional genes. Such cells may not survive as expanded clones. However,

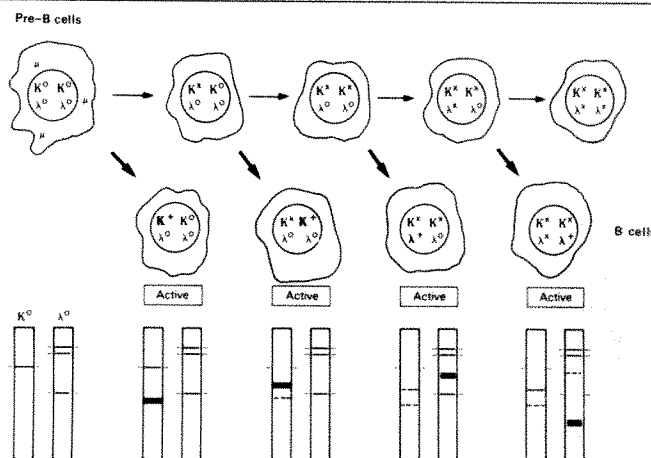


Fig. 6 Developmental scheme of ordered light-chain gene rearrangements in pre-B cells. The genotype of each intermediate cell is indicated within the nucleus, while the predicted arrangement of *Bam*HI(κ) or *Eco*RI(λ) fragments is shown diagrammatically below each active cell. The overall scheme of lymphocyte development is not shown in order to focus on those stages concerned with light-chain gene activation. The initial cell in this scheme represents a pre-B cell already committed to the expression of cytoplasmic μ heavy chain. The expected germ-line configuration of the κ (κ^0) and λ (λ^0) light-chain gene fragments is shown schematically below. Light-chain gene rearrangements then take place (either productively or aberrantly) in a sequence such that rearrangements progress generally from the κ to the λ light-chain genes. Gene rearrangements cease to occur as a result of the formation of an active light-chain gene and synthesis of a viable light-chain product as suggested by Alt *et al.*⁹. The genotypes and representative light-chain gene fragment configurations for each of the four light-chain-producing cell types are shown. Bold faced lines represent the active allele; a dotted line represents either a deleted or aberrantly rearranged allele. Genotype superscripts signify the following: 0, germ-line configuration; +, viable V-J joining event; X, aberrant rearrangement or deletion. Gene fragment patterns consistent with each active cell type have been noted in this study. In addition, the scheme predicts a population of B-cell precursors with non-rearranged or only one aberrantly rearranged gene. It also predicts a population of cells incapable of forming valid light-chains, having 'spent' their germ-line genes on invalid recombinations.

in a B cell in which a functional light-chain gene has been formed, it is reasonable to suggest that germ-line or aberrantly rearranged genes account for allelic exclusion. Perry *et al.*⁴ have emphasized that non-viable gene rearrangements may have a significant role in this process, whereas Joho and Weissman³ have interpreted their studies of κ gene arrangements in circulating lymphocytes so as to suggest that the excluded allele is most often in the germ-line configuration. Similar arguments can be made to account for isotype exclusion. The present study suggests that both aberrantly rearranged genes and genes that remain in the germ-line configuration are responsible for these types of exclusion.

The order of light-chain gene rearrangements

The fact that λ genes seem to remain in the germ-line configuration in κ -producing cells while κ genes are rearranged or (more often) deleted in λ -producing cells allows us to suggest a hierarchy of events involved in light-chain activation. If we simply assume that κ gene rearrangement usually precedes that of λ , and if we accept the useful proposal of Alt *et al.*⁹ that the appearance of a functional light-chain product precludes additional light-chain-gene rearrangements, we can place each of our B cells along the developmental scheme shown in Fig. 6. In this scheme it is proposed that a pre-B cell undergoes either a valid or a non-functional κ rearrangement. If the initial rearrangement is valid, the clone expands and the gene arrangement $\kappa^+/\kappa^0/\lambda^0/\lambda^0$ (where the superscripts + = valid, 0 = germ-line, X = aberrantly rearranged or deleted) is observed. If the initial κ rearrangement is non-functional, a second κ rearrangement occurs and, if valid, the clone expands and the arrangement $\kappa^X/\kappa^+/\lambda^0/\lambda^0$ is observed. If both κ alleles are aberrantly rearranged or deleted, the next rearrangement involves one of the λ alleles which, following the cascade diagram shown in Fig. 6, could go on to produce all the possible combinations of valid and non-functional genes shown in the figure. An intermediate population of pre-B cells which have no active light-chain genes

is predicted by this scheme, including a terminal cell type in which all light-chain genes are aberrantly rearranged or deleted. Each active configuration predicted by this scheme occurs among the B cells we have examined. On the other hand, if λ rearrangement were to precede that of κ or to occur at the same relative frequency as κ , then we should have seen λ genes rearranged in κ -producing cells. We have not observed this situation in any of the κ -producing cells examined. A sequential process of light-chain gene rearrangement, which seems to involve the frequent deletion of C_κ genes and a joining order that proceeds from κ to λ , can be thought of as an error-compensating cascade that finally creates an active gene while experimenting with the flexible joining rules that allow optimum diversification.

Why is the κ constant region deleted?

The arguments outlined above do not deal with the finding that the κ constant region genes are entirely deleted in 9 of the 10 λ chain-producing cells analysed. Such deletions obviously involve recombination sites not ordinarily involved in V/J recombination. For example, the deletion of C_κ may be the result of an aberrant recombination event between a V_κ gene and a recombination site which lies 3' to the κ constant gene. Alternatively, this deletion might represent secondary recombinational events that follow the formation of valid or invalid κ genes. But why has the gene disappeared with such a high frequency? One possibility is that this deletion is a normal step in

the progress of light-chain recombination, perhaps a step for the elimination of inactive κ genes. This seems unlikely in view of the fact that cells with active κ genes seem to tolerate the presence of invalid allelic recombinants. It is more likely that elimination of the κ gene is closely related to λ gene rearrangement, either as an effect or as a cause. The possibility that λ rearrangement is necessary for κ gene deletion seems doubtful in view of the several examples of κ deletion in κ chain-producing B cells in which no λ gene rearrangement has occurred (Fig. 1, cases 5 and 8). Instead, one might imagine the formation of a positive control element by the κ deletion event which causes the onset of λ gene rearrangements. Alternatively, a negative control element preventing λ gene rearrangement might be eliminated by the κ deletion event. Our single example (Fig. 4, case 10) of a retained (but rearranged) κ gene in a λ -producing B-cell line does not rule out the negative control argument as the effect might be sensitive to gene dosage or the κ gene rearrangement may itself have been sufficient to eliminate the control element. In addition, because this human cell line is uncloned, we must consider the possibility that the rearranged κ gene might reside in a subpopulation of B cells that is not involved in λ chain production. The striking observation that κ constant region genes are consistently deleted or rearranged in λ -expressing B cells may, among other possibilities, indicate a potential coupling between λ chain activation and κ gene deletion, a recombination event that might have regulatory implications.

Received 13 November 1980; accepted 14 January 1981.

- Melchers, F., von Boehmer, H. & Philips, R. A. *Transplant Rev.* **25**, 26–58 (1975).
- Seidman, J. G. & Leder, P. *Nature* **276**, 790–795 (1978).
- Joho, R. & Weissman, I. *Nature* **284**, 179–181 (1980).
- Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937–1941 (1980).
- Max, E. E., Seidman, J. G., Miller, H. & Leder, P. *Cell* **21**, 793–799 (1980).
- Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287**, 603–607 (1980).
- Seidman, J. G. & Leder, P. *Nature* **286**, 779–783 (1980).
- Leder, P. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
- Alt, F. W., Enea, V., Bothwell, A. L. & Baltimore, D. *Cell* **21**, 1–12 (1980).
- Hieter, P. A., Max, E. E., Seidman, J. G., Maizel, J. V. & Leder, P. *Cell* **22**, 197–207 (1980).
- Fett, J. & Deutsch, H. *Immunochemistry* **12**, 643 (1975).
- Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).
- Seidman, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6022–6026 (1980).
- Lenhard-Schuller, R., Hohn, B., Brack, C., Hiram, M. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4709–4713 (1978).
- Wilson, R., Miller, J. & Storb, U. *Biochemistry* **18**, 5013–5021 (1979).
- Steinmetz, M., Zachau, H. G. & Mach, B. *Nucleic Acids Res.* **6**, 3213–3229 (1979).
- Steinmetz, M. & Zachau, H. G. *Nucleic Acids Res.* **8**, 1693–1707 (1980).
- Wilson, R., Miller, J. & Storb, U. *Biochemistry* **18**, 5013–5021 (1979).
- Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3454–3460 (1979).
- Polisky, F., Edgell, M. H., Seidman, J. G. & Leder, P. *Analyt. Biochem.* **87**, 397–410 (1978).
- Alwine, J. C. *et al. Meth. Enzym.* **68**, 220–242 (1980).
- Rigby, P. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
- Sletten, K., Hannestad, K. & Harbol, M. *Scand. J. Immun.* **3**, 215–218 (1974).
- Kabat, E. A., Wu, T. T. & Bilofsky, H. *NIH Publ. No.* 80–2008, p. 10 (1979).

Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression

Christopher Coleclough, Robert P. Perry, Klaus Karjalainen & Martin Weigert

Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA

A study of the organization of light- and heavy-chain immunoglobulin genes in mouse splenic B cells, spleen-derived hybridomas and plasmacytomas has unequivocally demonstrated that aberrant rearrangements are common during normal B-cell development. The results support a probabilistic model for allelic exclusion of immunoglobulin gene expression.

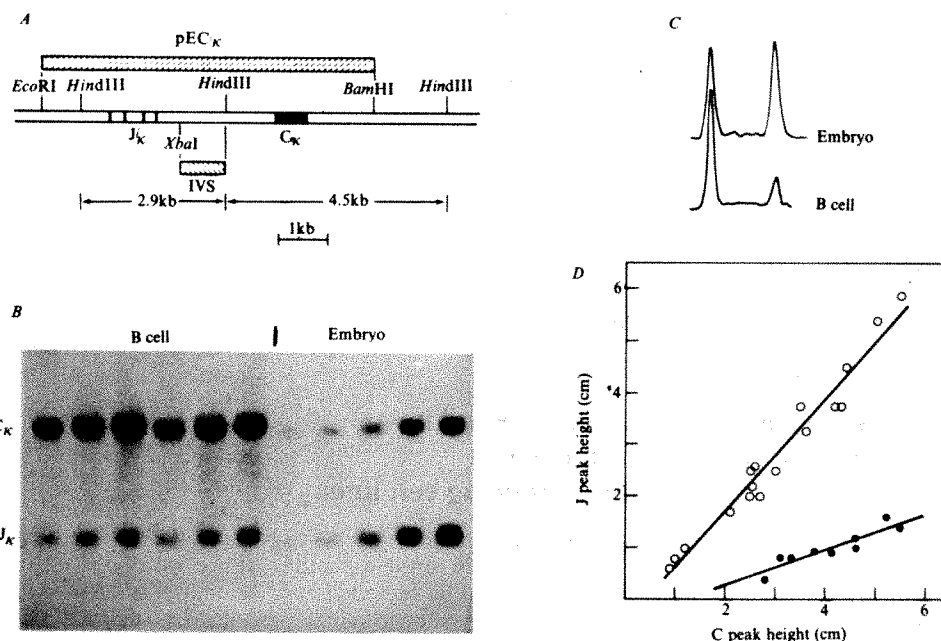
PHENOTYPIC expression of immunoglobulin genes depends on site-specific recombination events which bring the genetic elements encoding the immunoglobulin variable and constant regions within a single transcription unit. This is most easily appreciated in the mouse λ_1 light-chain gene system, in which an active gene is constructed by the precise fusion of the single V_{λ_1} gene with the single J_{λ_1} element residing 1.2 kilobases 5' of the C_{λ_1} gene¹. Although the mouse κ light-chain gene family is more complex, comprising some hundreds of V_κ genes and four functional J_κ elements, gene activation is essentially similar, requiring a single recombination event to fuse V_κ and J_κ ^{2,3}.

However, the construction of a functional μ heavy-chain gene seems to require two recombination events, which result in the fusion of three distinct genetic elements: V_H , D_H and J_H ^{4,5}.

Immunoglobulin gene expression exhibits allelic exclusion; that is, in each individual antibody-forming cell, only one of the pair of alleles at an immunoglobulin locus is expressed while the other is phenotypically silent. Similarly, the synthesis of κ and λ light chains in individual lymphoid cells is mutually exclusive. How is this achieved? Recently, we analysed the status of allelic κ genes in a large number of κ -secreting plasmacytomas⁶. In about half of the plasmacytomas examined, the phenotypically

Fig. 1 The extent of J_{κ} rearrangement in κ -bearing B cells. **A**, schematic map of the J_{κ} - C_{κ} region. Cloned hybridization probes are shown as hatched boxes. Plasmid pEC $_{\kappa}$ was constructed by insertion into pBR322 of the *EcoRI*-*Bam*HI fragment shown, which was derived from the 15.5-kilobase (kb) *EcoRI* fragment cloned in Charton 4A by Lenhard-Schuller *et al.*²³. IVS was isolated by *Hind*III-*Xba*I digestion of pEC $_{\kappa}$ and gel electrophoresis. Recombination involving the J_{κ} region will deplete the germ-line J_{κ} -containing *Hind*III fragment relative to the C_{κ} -containing fragment, and may be detected as a decrease in the J_{κ} / C_{κ} fragment ratio. Recombination 3' (right) of the *Hind*III site between J_{κ} and C_{κ} , however, is not detectable by this assay, as it depletes the germ-line J_{κ} and C_{κ} *Hind*III fragments to the same extent. **B**, a representative experiment. DNA samples prepared from 14-day BALB/c embryos²⁴, or from κ -bearing B cells were digested to completion with *Hind*III (Biolab), and various amounts (~0.5–8 μ g) electrophoresed through 0.8% neutral agarose gels, transferred to nitrocellulose paper²⁵ (Schleicher and Schuell), and hybridized to ³²P-labelled pEC $_{\kappa}$ plasmid DNA^{11,26}. Filters were exposed to X-ray film at

–70 °C, using an intensifying screen, for various periods; the one shown was exposed for ~10 h. The κ -bearing B cells were purified from BALB/c spleen cells. Suspensions of spleen cells were depleted of T cells by treatment with anti-Thy-1 and complement, and centrifuged through Ficoll-Hypaque (Pharmacia) to remove erythrocytes and dead cells. The B-cell-enriched preparation was labelled with rabbit antiserum directed against mouse κ constant region, counterstained with fluorescent-coupled goat anti-rabbit immunoglobulin and passed through the fluorescence-activated cell sorter (Becton-Dickinson, FACS II). **C**, densitometer tracings of representative tracks from Southern blot autoradiographs of embryo and B-cell DNA showing the marked reduction of the J_{κ} peak (right) in B-cell DNA. **D**, graphical estimation of the extent of J_{κ} rearrangement in B cells. Several exposures of the autoradiogram in **B** and another set of data from a similar experiment were scanned as in **C**. Comparison of multiple exposures of the same tracks showed that peak heights between 0.5 and 6 cm corresponded to the linear portion of the dose-response curve of the film. Data within this range were plotted as J_{κ} against C_{κ} ; each point is derived from an individual track. The least-squares best fit of the embryo data has a slope of 1.105; that of the B-cell data, 0.35. ○, Embryo DNA; ●, B-cell DNA.



silent allele was maintained in the germ-line context, whereas in the other half it had undergone some type of non-productive rearrangement. We concluded that imprecision intrinsic in the process of somatic recombination contributes significantly to the allelic exclusion of κ gene expression, and we proposed a model in which the status of the κ alleles is determined by the recombination frequency and the relative probabilities of productive as opposed to non-productive rearrangements^{6,7}. Our conclusion, which was supported by the results of other studies on the extent of J_{κ} ^{23,29,30} and J_{H} ³¹ rearrangement in myelomas and lymphomas, was based on the assumption that plasmacytomas are valid models of untransformed plasma cells. This assumption was questioned by Joho and Weissman⁸, who measured the extent of κ gene rearrangement in κ -expressing splenic B cells and concluded that non-productively rearranged κ genes are very rare or absent in normal lymphoid populations. However, their conclusion was based on a single determination and an arbitrary interpretation of their normalization data.

To settle this issue and to examine the general applicability of the probabilistic model we have made a detailed study of the extent of non-productive rearrangement of both light- and heavy-chain immunoglobulin genes in normal mouse splenic B cells, spleen cell-derived hybridomas and, for comparison, appropriate panels of plasmacytomas and B-cell lymphomas. Our results indicate that at least one out of three normal B cells contains a non-productively rearranged J_{κ} locus and that the silent J_{H} locus is rearranged in >80% of the B-cell population. We also provide evidence for extensive J_{κ} rearrangement in λ -producing cells in contrast to a paucity of J_{κ} rearrangement in κ expressors. These results demonstrate unequivocally that aberrant rearrangements are a common feature of normal B-cell development. Moreover, they support the idea that both class and allelic exclusion can be explained in terms of a probabilistic model of immunoglobulin gene rearrangement.

Non-productive rearrangement of κ genes occurs frequently in B cells

In monoclonal lymphoid lines, J_{κ} rearrangement is detected by observing a discrete J_{κ} -containing restriction fragment differing

from that found in embryo (germ-line) DNA. In contrast, one would not expect a discrete novel J_{κ} band in a Southern blot of DNA from the total splenic population of κ -expressing B cells because the many different V-J combinations would result in a diverse array of fragments of varying sizes. The extent of J_{κ} rearrangement in this population can, therefore, be calculated from the relative decrease in intensity of the embryo-type J_{κ} band. We needed a strategy to estimate precisely this decrease, using controls and quantitative measurements internal to the experiment. Figure 1A shows a map of the C_{κ} region. It can be seen that a *Hind*III digest of embryo DNA when probed on a Southern blot with a plasmid containing an appropriate *EcoRI*-*Bam*HI fragment, pEC $_{\kappa}$, reveals two major fragments. The larger of these, ~4.5 kilobases, contains the C_{κ} gene, while all the J_{κ} elements are contained within a fragment of ~2.9 kilobases. As these fragments overlap the pEC $_{\kappa}$ probe to about the same extent, their autoradiographic intensities should be roughly equivalent. That this is indeed the case was verified (Fig. 1B–D) by measurements of varying amounts of embryo DNA. As shown in Fig. 1D, a plot of the intensity of the J_{κ} band versus that of the C_{κ} band gives a straight line with a slope of ~1. For these determinations, autoradiographic exposure times were confined to a range in which the densitometric peak heights were directly proportional to the amount of radioactivity emitted. In contrast, when a similar set of measurements were made with the DNA of κ -expressing splenic B cells, the intensity of the J_{κ}

Table 1 Extent of J_{H} gene rearrangement in splenic B cells

Sample	C_{κ} gene dosage (μ g embryo DNA equivalent)	J_{H} gene dosage (μ g embryo DNA equivalent)	% J_{H} genes unrearranged	% B cells with both J_{H} genes rearranged*
B1	10.0	~0	~0	~100
B2a	8.0	0.5	6.25	87
B2b	7.5	0.7	9.3	81

The C_{κ} and J_{H} gene dosages were estimated from Fig. 5B.

* Assumes each cell has at least one rearranged J_{H} .

band was ~ 0.35 that of the C_κ band as judged by the slope of the J_κ versus C_κ plot. If all the B cells had contained one productively arranged and one unrearranged J_κ locus, the slope would have been 0.5. Our data demonstrate that this is not the case (confidence limit >0.95); rather they indicate that about one-third of the cells contain two rearranged J_κ loci, one of which would presumably be non-productive with respect to κ -chain production. This fraction is actually a minimum estimate because the above analysis is sensitive only to those rearrangements which result in a loss of the 2.9-kilobase J_κ fragment and retention of the 4.5-kilobase C_κ fragment, whereas some non-productive rearrangements could conceivably involve the loss of both fragments.

λ_1 genes are infrequently rearranged in κ -expressing cell lines but κ genes are often rearranged in λ -producing cell lines

We wished to assess whether the probabilistic model which we applied to allelic exclusion at κ loci is also valid for light-chain class exclusion. That is, can both κ and λ loci be regarded as elements competing for mutually exclusive expression in a single pool of B cells? If this is the case, then non-productive gene rearrangement would be evident in cell lines producing either κ or λ light chains.

Initially we determined the status of λ_1 genes in a panel of κ -producing plasmacytomas. Southern blots of *Eco*RI-digested embryo DNA probed with a λ_1 cDNA-containing plasmid exhibit four bands at 8.6, 4.8, 3.5 and 2.5 kilobases (Fig. 2). Brack *et al.*¹ have identified the largest three as containing C_{λ_1} , V_{λ_1} and V_{λ_2} , respectively. We do not know which hybridizing element is contained within the 2.5-kilobase band; possibly it is related to the λ_{III} locus⁹. Each of the four λ_1 -producing plasmacytomas examined showed a novel band at 7.4 kilobases, characteristic of the fused V_{λ_1} - C_{λ_1} gene¹. There is some evidence of additional λ gene rearrangement: Y5444 seems to lack an unrearranged allelic C_{λ_1} gene and J558 lacks an unrearranged

allelic V_{λ_1} gene. When DNA from 14 κ -producing plasmacytomas was similarly probed, all revealed the germ-line set of λ bands; one cell line, PC2960, showed one additional band that is different from the productive V_{λ_1} - C_{λ_1} hybrid. One λ gene has been rearranged non-productively in this cell line, but without further analysis, we cannot say which. Clearly, however, λ gene rearrangement is rare in κ -producing cells.

We performed the complementary experiment, determining the status of κ genes in λ -producing cells, at first with the four λ plasmacytomas used above. In no case could we find a germ-line-type κ gene (not shown). To extend our analysis, we decided to use λ -producing hybridomas, which can be produced in reasonable numbers and, moreover, can be assayed soon after induction, minimizing the time spent in serial passage. The use of cultured hybridoma lines also eliminates any possible contamination by DNA of non-lymphoid origin. Certain strains of mice, in response to primary immunization with nitrophenyl chicken globulin (NPCG), produce antibodies in which λ light chains are frequent¹⁰. We produced a large number of hybridomas from spleen cells after primary immunization of CB.17 mice with NPCG. After cloning and serological characterization, we analysed DNA from 12 λ -secreting hybridomas. DNA was digested with *Hind*III and processed as above, this time probing with the *Hind*III-*Xba*I fragment, IVS. As shown in Fig. 1A, embryo DNA yields a 2.9-kilobase fragment that hybridizes with this probe; rearrangement at or close to any J_κ will substitute this with a novel fragment. Our results are shown in Fig. 3. Ag8.653, the myeloma partner used to produce hybridomas, has a strong band at ~ 6.5 kilobases, which is maintained in all the hybridomas. Clearly, rearrangement at J_κ is frequent in these hybridomas. Four hybridomas, *d*, *e*, *f* and *g*, showed bands similar in mobility to that characteristic of unrearranged C_κ genes. A second analysis, including an embryo marker, shows that only the fragment from hybridoma *f* is actually coincident with the embryo-type fragment (Fig. 3B). To verify that hybridoma *f* does indeed contain a non-rearranged κ gene, we performed a similar analysis with DNA that was digested with *Bam*RI and *Bam*HI plus *Eco*RI. Hybridoma *f* continues to

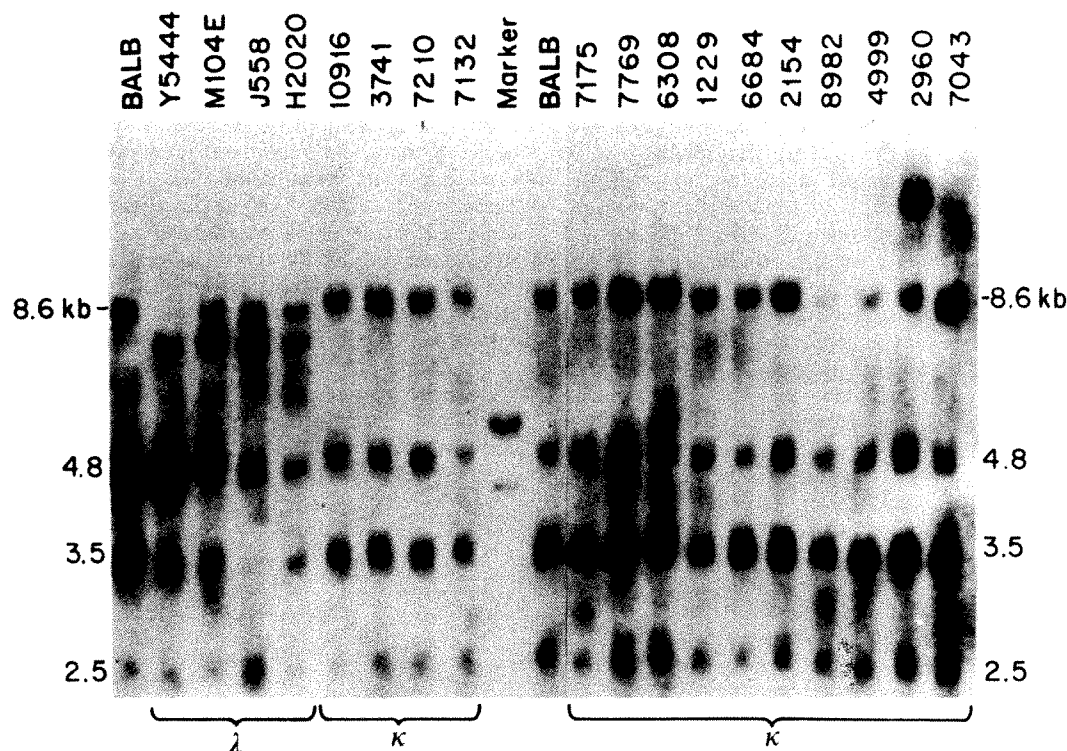
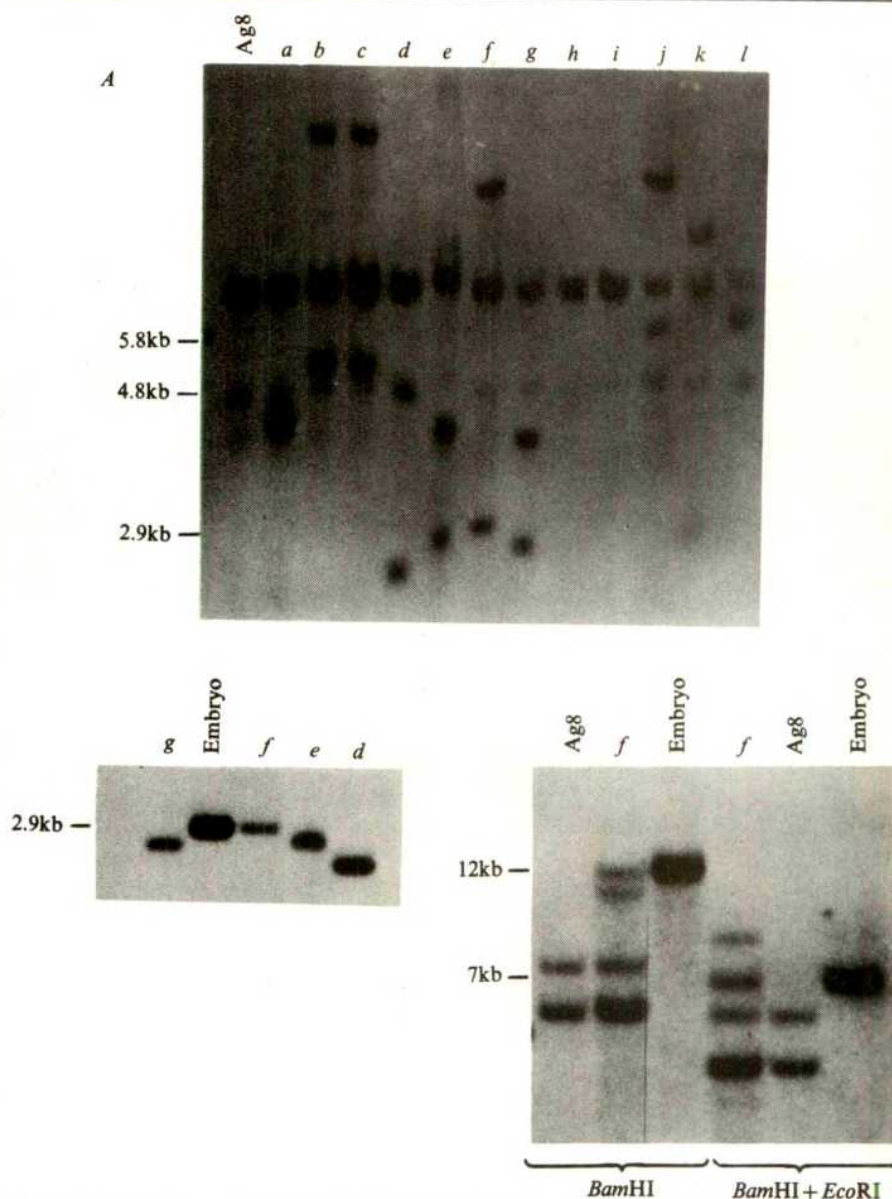


Fig. 2 Southern blot analysis of the λ loci in a series of λ - and κ -secreting plasmacytomas. Aliquots (10 μ g) of DNA from the sources indicated were digested with *Eco*RI, electrophoresed through a 0.7% agarose gel, transferred to nitrocellulose and hybridized with ^{32} P-labelled plasmid B1, which contains V_{λ_1} and C_{λ_1} sequences derived from H2020 mRNA¹. Tracks marked BALB contained 10 μ g 14-day BALB/c embryo DNA. kb, Kilobase.

Fig. 3 J_{κ} rearrangement in λ -producing hybridomas. Hybridomas specific for NPCG were produced as described elsewhere¹⁰ by fusion of immune spleen cells with Ag8.653 cells²⁷. After selection for drug-resistance phenotype and antigen binding specificity, those hybridomas secreting λ light chains were revealed by radioimmunoassay. **A**, DNA from the myeloma parent cell line Ag8.653 and from each of 12 individual cloned λ -producing hybridomas (*a-l*) was digested with *Hind*III, electrophoresed through 0.7% agarose and probed with ³²P-labelled IVS (see Fig. 1A). Ag8.653 DNA shows strong hybridization to a fragment ~6.5-kilobases long, and this is maintained in all hybridomas. The weak band at 4.5 kilobases is due to slight contamination of the electrophoretically purified IVS probe with a C_{κ} fragment, which hybridizes to the 4.5-kilobase C_{κ} *Hind*III fragment and thus is maintained in all cell lines. The J_{κ} region in its germ-line context is revealed as a 2.9-kilobase fragment. **B**, four hybridomas, *d-g*, which exhibited J_{κ} *Hind*III fragments with mobilities close to that characteristic of the germ-line J_{κ} region, were re-assessed relative to BALB/c embryo DNA. *Hind*III digests were analysed by gel electrophoresis and Southern blotting as above. **C**, hybridoma *f* was further analysed to confirm that it contained a non-rearranged J_{κ} region; *f* has a fragment with identical mobility to that of the germ-line fragment after both *Bam*HI digestion and *Bam*HI-*Eco*RI double digestion.



display a fragment indistinguishable in size from the embryo-type κ gene fragment (Fig. 3C).

Thus, the prevalence of rearranged over non-rearranged κ genes in λ -producing lymphoid lines is high but not absolute. Of the 15 distinct J_{κ} gene contexts (counting *b* and *c* as one clone; see below) discovered in these 12 hybridomas, one is germ line and 14 are rearranged. Three of the hybridomas show only one J_{κ} fragment; two show no hybridization other than to the Ag8 fragment. This could be due either to chromosome segregation or to the specific deletion of this region. Interestingly, two clones, *b* and *c*, show identical patterns of J_{κ} context, and are also identical in heavy-chain gene context (not shown). We are certain that these are independent clones; thus they probably derive from the same B cell, clonally expanded *in vivo*.

J_H rearrangement usually occurs in both heavy-chain alleles in plasmacytomas and normal B cells

We set out to evaluate the incidence of rearrangement in the silent J_H locus by examining the context of the J_H cluster in established lymphoid lines and splenic B cells. The J_H elements are contained in an *Eco*RI fragment ~6.2 kilobases long (Fig. 4A). Recombination within or near the J_H cluster should produce a novel *Eco*RI fragment which could be revealed by hybridization of an appropriate recombinant plasmid, pJ11. An application of this analysis to a series of plasmacytomas is shown

in Fig. 4B. As all the plasmacytomas synthesize a heavy chain, at least one novel *Eco*RI fragment is expected in each case. In fact, most, 15 out of 20, show two distinct novel *Eco*RI fragments, as do the two lymphomas 38C and 70Z, demonstrating frequent rearrangement of the silent allelic J_H cluster. All plasmacytomas which were grown as subcutaneous tumours show some, mostly faint, hybridization to an embryo-type fragment which presumably derives from contaminating host cells, as noted before^{6,11}. Rearrangement of both J_H alleles was also observed by Cory and co-workers³¹ in a similar analysis of a myeloma and several lymphoma lines. Three examples, PC2960, PC7043 and PC6308, show strong hybridization to an embryo-type fragment and only a single novel fragment. We have further analysed PC7043 and PC6308 using an *Eco*RI-*Sac*I digest, which yields a J_H fragment of ~3 kilobases. Both plasmacytomas yield a J_H band coincident with that of the embryo type (data not shown). Thus, we note that, although rearrangement of the silent allelic J_H cluster is frequent in lymphoid lines, it is not a universal occurrence.

In our first set of experiments with splenic B cells purified on a fluorescence activated cell sorter (FACS), we noted that, whereas *Eco*RI digestion of B-cell DNA yields a distinct C_{μ} fragment, no discrete band is obtained using *Kpn*I¹². Because the *Kpn*I site upstream of C_{μ} is 5' distal to the J_H cluster, this result suggested relatively frequent J_H recombination in B cells. Here we refine and quantify those observations.

As discussed above for J_{κ} rearrangement, the extent of J_H rearrangement in the splenic population of surface IgM-positive

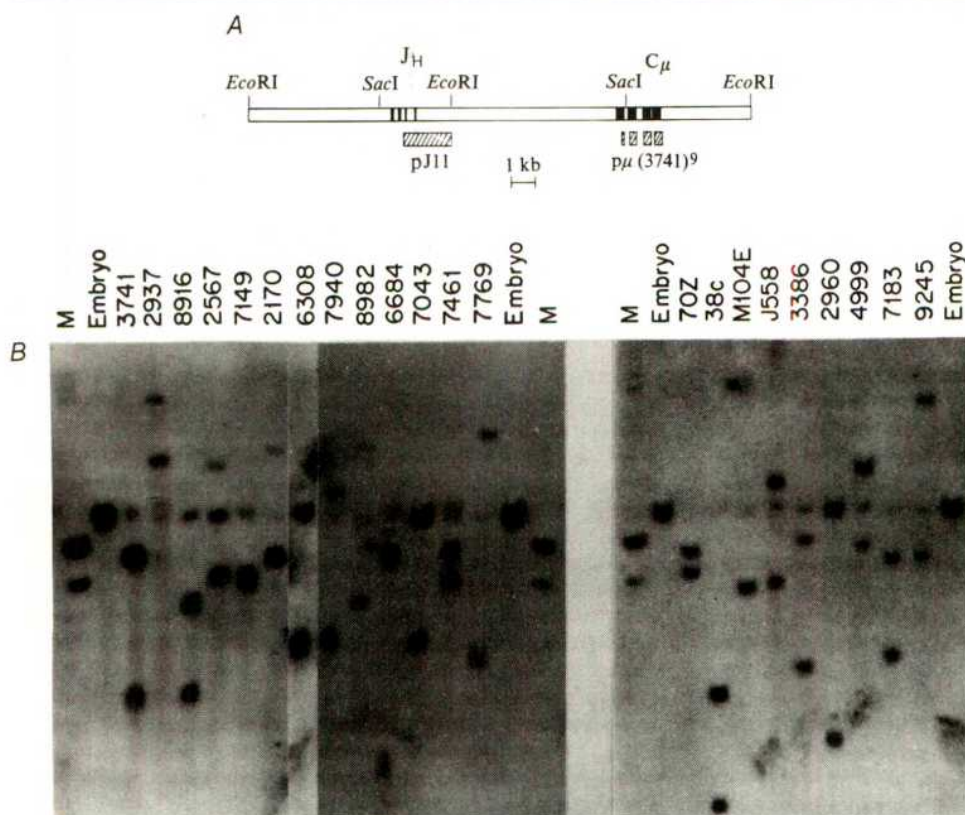


Fig. 4 J_H gene rearrangement in lymphoid lines. **A**, restriction map of C_μ gene and J_H region. Black areas indicate protein coding elements. Hatched boxes indicate sequences represented in the C_μ probe: the cDNA plasmid pμ(3741)⁹ and the J_H probe: the genomic DNA plasmid pJ11¹³. The map was compiled from refs 13 and 28 and our own observations. Not all *SacI* sites are shown. **B**, 10 μg of DNA prepared from each indicated source was digested with *EcoRI* and electrophoresed through 0.75% agarose, transferred to nitrocellulose and hybridized with ³²P-labelled pJ11. All cell lines are from NZB mice, except M104E and J558, which are from BALB/c mice, 38C from a C3H mouse, and 70Z which derives from a (C75BL/6 × DBA/2) F₁ hybrid. The mobility of non-rearranged J_H restriction fragments from these strains is identical, thus all tracks are comparable with the BALB/c embryo tracks. Tracks labelled M contained 1 μg *EcoRI*-digested λ phage DNA and 3 pg each of *EcoRI*-treated, and *EcoRI* plus *SacI*-digested pMB9 DNA. The latter species are revealed by hybridization with recombinant plasmid probes and appear on the autoradiograph as 5.8- and 4.8-kilobase fragments.

B cells must be calculated from the relative decrease in intensity of the embryo-type J_H band. We again used hybridization to the constant-region gene fragment to control for and quantify the extent of hybridization to the neighbouring J-region fragment. We wished to separate the C_μ gene from its labile 5' flanking region¹³ to avoid the possibility that any rearrangement in this region could lead to an underestimation of C_μ gene dosage. This was achieved using *SacI*, which cleaves in the intron between $C_{\mu 1}$ and $C_{\mu 2}$ exons. A *SacI*-*EcoRI* double digestion yields a 5.1-kilobase C_μ fragment and a 3-kilobase J_H fragment from embryo DNA (Fig. 4A).

Our analysis of splenic B cells is shown in Fig. 5A. Two different samples of B-cell DNA were analysed, one in duplicate, loading ~6 to 8 μg of DNA per track, together with two sets of a titrated series of embryo DNA. B-cell DNA shows a striking deficit of the J_H band relative to embryo DNA, indicating extensive rearrangement of the J_H region. One sample (B1) shows no distinct J_H band at all, whereas the other duplicate samples (B2a, b) show very faint bands over an elevated background.

Because of the large difference in intensity of the J_H and C_μ bands it was not possible to select a range of autoradiographic exposures for which both bands were on the linear portions of their dose-response curves. Therefore, we had to use a different method of calculating the extent of J_H rearrangement from that used for the J_κ measurements. By varying the exposure time of the autoradiograph and scanning the tracks on a densitometer, we obtained calibration curves relating C_μ of J_H peak height to the input amount of embryo DNA (Fig. 5B). By reference to these curves we used the measured peak heights to estimate the apparent C_μ gene and J_H gene dosage as equivalent μg total embryo DNA input (Table 1). The extent to which apparent J_H gene dosage is less than apparent C_μ gene dosage in B-cell DNA equals the extent of J_H rearrangement. Assuming that all IgM-expressing B cells must have one rearranged J_H region, we calculated the per cent of cells in each sample which have rearranged their silent J_H allele as: B1 ~100%, B2a 87% and B2b 81%. The persistence of a small proportion (~15%) of cells with an unrearranged J_H locus is consistent with the results of the plasmacytoma survey described above.

Discussion

These results clearly demonstrate that the aberrant rearrangements previously observed in transformed lymphoid cells also occur in normal B cells, and consequently that the recombination process responsible for the formation of functional immunoglobulin genes is error prone. Thus, when recombination events occur during B-lymphocyte maturation, a germ-line locus (Ig⁰) may be rearranged to form either a functional (Ig⁺) or a non-productive (Ig⁻) immunoglobulin gene. Detailed analyses of κ genes have revealed several types of non-productive rearrangements: those involving recombination at a pseudo J site, which lacks a proper RNA processing signal^{6,14,15}, those involving joining errors, which create translational stop codons^{16,17}, and those which extinguish transcription^{6,7}, possibly due to the selection of a non-functional (pseudo?) V-region segment.

We have suggested that the status of immunoglobulin alleles depends principally on their recombination frequency, which may vary over the course of B-cell maturation, and the relative probability of productive as opposed to non-productive rearrangements⁷. With appropriate values of these parameters, this probabilistic model should predict the ratio of Ig⁰/Ig⁺ to Ig⁻/Ig⁺ genotypes in a particular population of lymphoid cells. Our previous studies of κ -expressing plasmacytomas indicated a 1:1 ratio of κ^0/κ^+ to κ^-/κ^+ genotypes, whereas our current measurements of κ -expressing splenic B cells indicates a ratio of ~2:1. If this discrepancy actually reflects a difference between plasma cells and B cells, it would suggest that κ rearrangement can still occur during clonal expansion of B cells, although at a greatly diminished frequency compared with pre-B cells. Our estimate of 33% non-productive rearrangement in B cells is in marked contrast to the very low level claimed by Joho and Weissman⁸. However, when their single measurement was normalized with their rDNA data, we calculated 26% non-productive rearrangement, which is not significantly different from our more accurately determined value.

The high frequency of κ rearrangement in λ expressors and the rarity of λ rearrangement in κ expressors suggest that the recombination frequency is in part determined by the total

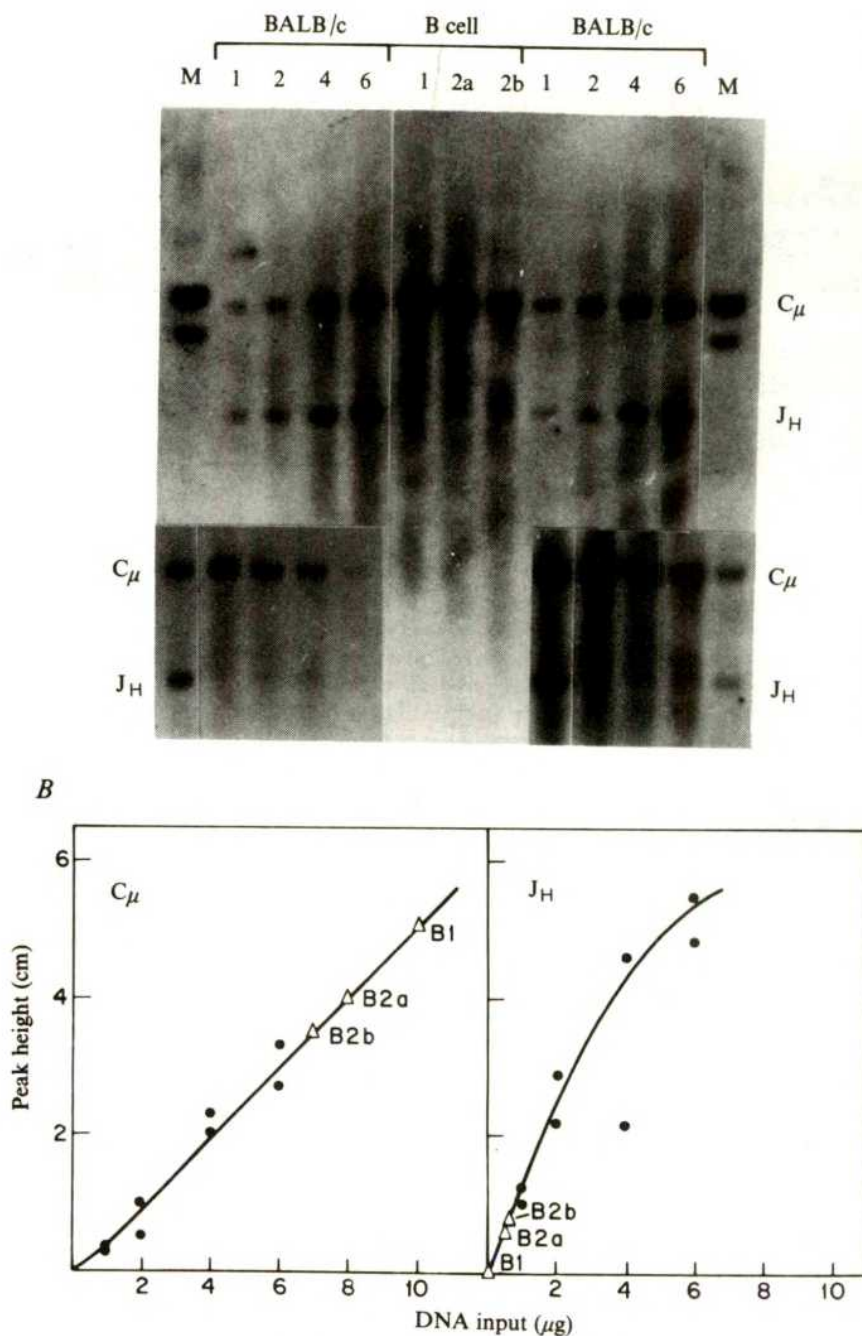
number of V and J elements that can participate in recombination, which in the mouse would greatly favour κ over λ . Furthermore, the fact that a κ^0 locus is occasionally retained by a λ expressor and a λ^- locus is occasionally present in a κ expressor is consistent with a probabilistic model which views the genes for both classes of light chain as equivalent candidates for productive or non-productive rearrangement. This fact is not easily accommodated by models which postulate a strictly sequential sampling of the κ and λ gene pools¹⁸.

If we consider the formation of a productive light-chain gene as the prime determinant of the developmental transition from pre-B to B cell¹⁹, then the composition of κ genotypes in λ -expressing spleen cells should reflect the steady-state distribution in pre-B cells at the time of this transition. Our results indicate a high proportion of κ^-/κ^- genotypes in these pre-B cells. For this to be compatible with a 2:1 ratio of κ^0/κ^+ to κ^-/κ^+ genotypes in B cells we need to postulate that the recombination frequency is greatly reduced after the formation of a productive light-chain locus. Such a supposition seems reasonable as the production of light chains allows the cell to display surface immunoglobulin²⁰ which presumably triggers

gross morphological and biochemical changes associated with the pre-B to B cell transition. This effective stabilization of genotypes by a productive rearrangement would effectively ensure the allelic exclusion of light-chain gene expression by diminishing the likelihood of $\kappa^0/\kappa^+ \rightarrow \kappa^+/\kappa^+$ transitions. A similar suggestion was recently made by Alt *et al.*¹⁸.

Our analysis of the heavy-chain (J_H) locus indicates a high frequency of non-productive rearrangement, with >80% of splenic B cells having μ^-/μ^+ genotypes. Yet the fact that a small but definite subpopulation have μ^0/μ^+ genotypes is consistent with the probabilistic nature of the recombination process. The difference in frequency of non-productive rearrangements between heavy- and light-chain genes may be related to the more complex structure of the heavy-chain locus, which requires at least two recombinational events (V_H to D_H and D_H to J_H) to produce a functional genetic unit. Lack of function may be due to an error in the D_H - J_H recombination, an error in the accompanying V_H - D_H fusion, or even the lack of a V_H - D_H recombination. In the latter case, the non-productive locus might represent an intermediate which could eventually undergo further rearrangement.

Fig. 5 J_H gene rearrangement in splenic B cells. A, B-cell-enriched fraction prepared as described in Fig. 1 legend was stained with fluorescein-coupled rabbit anti-mouse IgM and the surface IgM-positive fraction purified by FACS. The proportion of IgM-bearing cells in the final cell sample was estimated to be >98%. DNA was prepared from these cells and its concentration roughly estimated by visual inspection of ethidium bromide fluorescence of an aliquot compared with that of DNA of known concentration. B-cell DNA was digested sequentially with *SacI* and *EcoRI* and an estimated 6–8 μ g electrophoresed through 0.8% agarose. Two different samples of B-cell DNA were analysed, one of them in duplicate (B1; B2a, B2b). A titrated series of BALB/c embryo DNA was similarly analysed, in duplicate, loading 1, 2, 4 and 6 μ g as shown. Hybridization of the Southern blot to a mixture of ³²P-labelled plasmids p μ (3741)⁹ and pJ11 revealed the C_μ gene and the J_H region as fragments of 5.1- and 3.0-kilobases, respectively. Three exposures (20, 40 and 100 h) of the autoradiograph were developed. The entire 40-h exposure is shown, with the central five tracks (the three B-cell samples flanked on the left by 6 μ g and on the right by 1 μ g embryo DNA) of the 20- and 100-h autoradiographs shown as inserts on the lower left and lower right, respectively. B, the autoradiographs were scanned on a densitometer and peak heights measured. Peak heights of the C_μ fragment in embryo DNA from the 20-h autoradiograph were plotted against μ g input DNA to give a calibration curve linear at large DNA inputs, which we could use to estimate B-cell DNA input from C_μ peak height as shown. The curve relating embryo DNA input to J_H peak height on the 100-h autoradiograph is linear at small DNA inputs, and was used to estimate B-cell J_H gene dosage. ●, Embryo DNA; △, B-cell DNA.



A higher intrinsic recombinational frequency for heavy-chain compared with light-chain genes could explain why rearrangement of heavy-chain genes precedes that of light-chain genes during B-cell ontogeny^{21,22}. The fact that a relatively high proportion of the heavy-chain gene rearrangements are non-productive could account for the allelic exclusion of heavy-chain expression, for it would predict that most cells in which a $\mu^0 \rightarrow \mu^+$ transition occurs would already have experienced a $\mu^0 \rightarrow \mu^-$ on the other allele. This explanation is compatible with the idea that both heavy- and light-chain genes use the same recombinational machinery^{4,5}, and that this machinery remains essentially intact on formation of a functional heavy-chain gene. This seems reasonable because there are apparently no gross phenotypic changes in μ -expressing pre-B cells before the creation of a functional light-chain gene. Thus, allelic exclusion

of immunoglobulin gene expression is viewed as being due principally to recombinational errors in the case of heavy chains, and to a combination of errors and phenotypic alteration in the case of light chains.

We thank K. Marcu for the pJ11 plasmid, R. Riblet and J. Brill for many DNA samples, J. Haimovich for lymphoma DNA, B. Subbarao for fluorescent antibody and R. Overton for assistance with use of the FACS. The research was supported by a NSF grant, NIH institutional grants, a fellowship to K.K. from the EMBO and an appropriation from the Commonwealth of Pennsylvania.

Note added in proof: After submission of this article another observation of frequent rearrangement of the non-expressed J_H allele in normal B cells was reported³².

Received 16 September 1980; accepted 4 February 1981.

- Brack, C., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15**, 1-14 (1978).
- Max, E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
- Sakano, H., Huppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
- Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981-992 (1980).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
- Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937-1941 (1980).
- Perry, R. P., Coleclough, C. & Weigert, M. *Cold Spring Harb. Symp. quant. Biol.* (in the press).
- Joho, R. & Weissman, I. *Nature* **284**, 179-181 (1980).
- Azuma, T., Steiner, L. A. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **78**, 569-573 (1981).
- Karjalainen, K., Bang, B. & Mäkelä, O. *J. Immun.* **125**, 313-317 (1980).
- Coleclough, C., Cooper, D. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1422-1426 (1980).
- Hurwitz, J. L., Coleclough, C. & Cebra, J. J. *Cell* **22**, 349-359 (1980).
- Marcu, K., Banerji, J., Pennicavage, N. A., Lang, R. & Arnheim, N. *Cell* **22**, 187-196 (1980).
- Choi, E., Kuehl, M. & Wall, R. *Nature* **286**, 776-779 (1980).

- Seidman, J. G. & Leder, P. *Nature* **286**, 779-783 (1980).
- Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287**, 603-607 (1980).
- Max, E. E., Seidman, J. G. & Leder, P. *Cell* **21**, 793-794 (1980).
- Alt, F. W., Enea, V., Bothwell, A. L. M. & Baltimore, D. *Cell* **21**, 1-12 (1980).
- Levitt, D. & Cooper, M. D. *Cell* **19**, 617-625 (1980).
- Paige, C. J., Kincade, P. W. & Ralph, P. J. *Immun.* **121**, 641-647 (1978).
- Maki, R., Kearney, J., Paige, C. & Tonegawa, S. *Science* **209**, 1366-1369 (1980).
- Perry, R. P., Kelley, D. E., Coleclough, C. & Kearney, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 247-251 (1981).
- Lenhard-Schuller, R., Hohn, B., Brack, C., Hiram, M. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4709-4713 (1978).
- Perry, R. P., Kelley, D. E., Schibler, U., Huebner, K. & Croce, C. M. *J. cell. Physiol.* **98**, 553-560 (1979).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237-251 (1977).
- Kearney, J., Radbruch, A., Liesgand, B. & Rajewsky, K. *J. Immun.* **123**, 1548-1550 (1979).
- Davis, M. M. *et al. Nature* **283**, 733-739 (1980).
- Wilson, R., Miller, J. & Storb, U. *Biochemistry* **18**, 5013-5021 (1979).
- Steinmetz, M., Zachau, H. G. & Mach, B. *Nucleic Acids Res.* **6**, 3213-3229 (1979).
- Cory, S., Adams, J. M. & Kemp, D. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4943-4947 (1980).
- Nottenburg, C. & Weissman, I. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 484-488 (1981).

LETTERS

Cyclotron and annihilation lines in γ -ray bursts

E. P. Mazets, S. V. Golenetskii, R. L. Aptekar',
Yu. A. Gur'yan & V. N. Il'inski

A. F. Ioffe Physico-Technical Institute, Academy of Sciences of the USSR, 194021 Leningrad, USSR

Venera 11 and Venera 12 have detected emission and absorption lines in the energy spectra of many γ -ray bursts. The similarity between the spectral features observed during different events reveals the similarity in physical conditions and processes which occur within the burst sources. These lines are the most convincing evidence that the γ -burst sources contain neutron stars with a strong magnetic field. We present here results of a comprehensive analysis that included weaker bursts, which indicate that the presence of lines in the γ -burst spectra, in particular of absorption lines in the region 30-70 keV is a characteristic rather than an exceptional feature.

The investigation of lines in cosmic X-ray and γ -radiation is important, and has been reviewed elsewhere^{1,2}. γ -ray line radiation is characteristic of nuclear processes, however, in certain conditions, lines can be produced in the astrophysical plasma by radiative processes^{3,4}. Lines in the plasma radiation can be closely associated with the presence of superstrong gravitational and magnetic fields in sources such as degenerate dwarfs and neutron stars. The magnetic bremsstrahlung at the cyclotron frequency $\omega_H = eH/mc$ at fields $H > 10^{12}$ Oe already lies in the hard X-ray range. Pair production in the hot magnetized plasma may be a powerful source of positrons. The positron annihilation radiation produced near the surface of a neutron star should undergo a substantial gravitational redshift. Observations of these effects could provide a means for reliable determination of the gravitational and magnetic fields in the sources.

Cyclotron radiation in the X-ray range was first observed⁵ as the 58-keV line in the spectrum of Her X-1; the 73-keV line in the Crab spectrum may have the same nature⁶. It is thought that the redshifted annihilation line was observed⁷ among other lines in a transient event on 10 June 1974 and in the emission from the Crab nebula region⁸. However, not all of these pioneering

Table 1 Emission lines in the γ -ray burst spectra

Date of event	T_0 (UT)	Full energy flux S (erg cm ⁻²)	$\bar{E}$ (keV)	F (cm ⁻² s ⁻¹)	Energy in line S_e (erg cm ⁻²)	S_e/S
18 September 1978	19.49.11	2.8×10^{-5} (2.5×10^{-5})	400 ± 50	0.43 ± 0.10	$(2.1 \pm 0.5) \times 10^{-6}$	0.07 (0.08)
4 November 1978	16.17.53	2.6×10^{-4} (1.4×10^{-4})	400 ± 50	1.22 ± 0.19	$(5.7 \pm 0.9) \times 10^{-6}$	0.02 (0.04)
19 November 1978	09.28.33	5.5×10^{-4} (5.0×10^{-4})	330-850	15.0 ± 1.6	$(4.3 \pm 0.5) \times 10^{-5}$	0.08 (0.09)
16 January 1979	08.59.18	2.3×10^{-5}	420 ± 50	0.19 ± 0.07	$(1.9 \pm 0.7) \times 10^{-6}$	0.08
5 March 1979	15.51.39	1.3×10^{-3} (3.0×10^{-4})	430 ± 30	73 ± 15	$(1.0 \pm 0.2) \times 10^{-5}$	0.008 (0.03)
18 April 1979	07.44.41	6.5×10^{-5}	450 ± 40	0.26 ± 0.08	$(2.7 \pm 0.8) \times 10^{-6}$	0.04
22 May 1979	21.50.01	7.4×10^{-6}	45 ± 5	0.9 ± 0.1	$(7.9 \pm 0.8) \times 10^{-7}$	0.11
22 June 1979	00.41.30	7.2×10^{-5}	460 ± 40	0.12 ± 0.04	$(3.5 \pm 1.0) \times 10^{-6}$	0.05

Fig. 1 The energy spectra of γ -bursts. *a*, 7 March 1979; *b*, *c*, 1 November 1979 (illustrating the cyclotron absorption line evolution in time); the spectra shown were obtained in the first 8 s (*b*) and subsequent 24 s (*c*) of the burst; *d*, 12 June 1979; *e*, 12 April 1979 (initial stage); *f*, 10 June 1979.

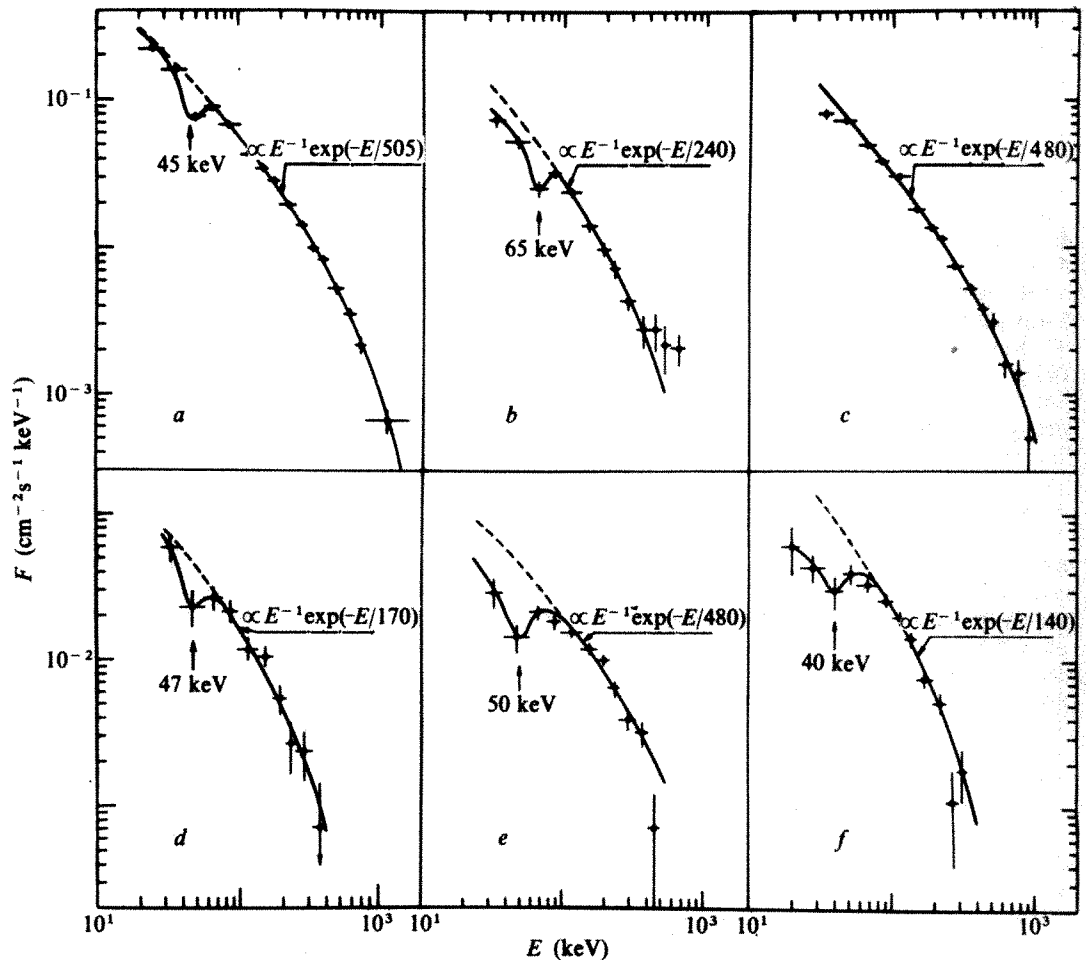


Table 2 Absorption lines in the γ -ray burst spectra

Date of event	T_0 (UT)	Full energy flux S (erg cm^{-2})	$\bar{E}$ (keV)	Relative depth h	Equivalent width W (keV)	Fraction of energy absor- bed S_a (erg cm^{-2})	S_a/S
14 September 1978	16.42.12	1.7×10^{-5}	65 ± 10	Wide absorption band			
17 November 1978	13.46.42	1.2×10^{-5} (8.0×10^{-6})	52 ± 8	0.70	28 ± 5	$(6.2 \pm 1.0) \times 10^{-7}$	0.05 (0.08)
21 November 1978	01.34.14	9.0×10^{-5} (1.5×10^{-5})	70 ± 10	0.36	23 ± 5	$(8.0 \pm 1.5) \times 10^{-7}$	0.009 (0.05)
7 March 1979	22.15.34	1.9×10^{-4}	45 ± 5	0.43	7.0 ± 0.6	$(2.3 \pm 0.2) \times 10^{-6}$	0.012
23 March 1979	23.39.11	3.4×10^{-5} (1.1×10^{-5})	66 ± 10	0.43	30 ± 5	$(1.1 \pm 0.2) \times 10^{-6}$	0.03 (0.10)
25 March 1979	13.41.12	4.8×10^{-5} (1.9×10^{-5})	50 ± 10	0.32	12.0 ± 1.3	$(5.6 \pm 0.6) \times 10^{-7}$	0.012 (0.03)
29 March 1979	22.28.42	6.5×10^{-5} (3.7×10^{-5})	50 ± 8	0.35	3.9 ± 0.7	$(2.6 \pm 0.5) \times 10^{-7}$	0.004 (0.007)
2 April 1979	09.41.45	3.7×10^{-5}	55 ± 15	Wide absorption band			
12 April 1979	22.04.42	2.2×10^{-5} (1.7×10^{-5})	50 ± 8	0.62	27 ± 2	$(1.8 \pm 0.2) \times 10^{-6}$	0.08 (0.10)
24 May 1979	03.56.26	1.5×10^{-5}	27 ± 5	0.31	3.4 ± 0.9	$(1.9 \pm 0.5) \times 10^{-7}$	0.013
29 May 1979	20.23.01	7.4×10^{-6}	55 ± 10	0.46	20 ± 5	$(3.5 \pm 0.9) \times 10^{-7}$	0.05
10 June 1979	05.50.04	6.1×10^{-6}	40 ± 5	Wide absorption band			
12 June 1979	04.42.54	3.0×10^{-6}	47 ± 7	0.49	13 ± 4	$(2.5 \pm 0.8) \times 10^{-7}$	0.08
22 June 1979	00.41.30	7.2×10^{-5}	50 ± 8	0.48	9.5 ± 0.8	$(1.3 \pm 0.2) \times 10^{-6}$	0.018
31 July 1979	10.49.55	8.1×10^{-5} (3.1×10^{-5})	55 ± 10	0.40	21 ± 2	$(9.3 \pm 1.1) \times 10^{-7}$	0.011 (0.03)
6 October 1979	08.10.26	7.3×10^{-6} (2.2×10^{-6})	62 ± 8	0.52	28 ± 4	$(4.0 \pm 0.6) \times 10^{-7}$	0.055 (0.18)
31 October 1979	09.18.33	6.5×10^{-6} (5.4×10^{-6})	46 ± 8	Wide absorption band			
1 November 1979	17.36.27	2.1×10^{-4} (3.4×10^{-5})	65 ± 10	0.47	18 ± 2	$(5.9 \pm 0.8) \times 10^{-7}$	0.003 (0.017)
9 November 1979	07.56.32	6.8×10^{-6}	64 ± 10	0.31	26 ± 3	$(8.4 \pm 0.9) \times 10^{-7}$	0.12
20 December 1979	17.30.34	1.0×10^{-4} (5.7×10^{-5})	63 ± 10	0.22	18 ± 1	$(1.6 \pm 0.1) \times 10^{-6}$	0.016 (0.028)

results are reliable²; in particular, the spectral features from the Crab nebula have not shown up in other observations therefore independent confirmation is needed.

Energy spectrum studies and the search for lines are also important in understanding the nature of the γ -bursts. In the Konus experiment carried out on Venera 11 and Venera 12 spacecraft during 1978–80 (ref. 9) energy spectra of ~ 150 γ -ray bursts in the 30 keV–2 MeV range were obtained, and the very first observations¹⁰ revealed lines in the spectra of some bursts. A redshifted annihilation line was clearly observed at 430 keV in the spectrum of the 5 March 1979 burst produced by a flaring X-ray pulsar¹¹. A broad emission feature was detected at 300–800 keV in the spectrum of a strong γ -burst on 19 November 1978 (ref. 9). A germanium detector on ISEE 3 revealed this as two lines at ~ 420 and 740 keV which were interpreted as the strongly redshifted 511-keV annihilation line and 847-keV iron line, respectively¹². Our subsequent measurements showed the presence of emission and absorption lines in several strong γ -bursts¹³.

The most typical examples of γ -burst spectra containing lines are shown in Figs 1–3. (More information on these events is given in refs 14, 15.) Eight consecutive measurements of the energy spectrum were made during each event, each lasting 4 s. Thus between 1 and 8 spectra were obtained for each burst, depending on its duration. The statistical accuracy of an individual spectrum of a weak event may be insufficient for line resolution, in which case, individual spectra are added together. This summation can also be used if no noticeable evolution of the spectrum is observed in the course of a burst.

Figure 1a shows the spectrum of a long intense event on 7 March 1979. Here the spectral distribution in the continuum, just as in most of the other bursts, fits the $dN/dE \propto$

$E^{-1} \exp(-E/kT)$ law typical for radiation from a hot, optically-thin plasma. An absorption line can be clearly seen at ~ 45 keV remaining visible in all the measured spectra. The line is assumed to result from absorption of the burst radiation at the cyclotron frequency ω_H in the outer, colder regions of the plasma in a strong magnetic field. The resonance is associated with the change of state of the electron in a magnetic field, its transition to the next Landau level involving the absorption of an energy $\Delta E = \hbar\omega_H = \hbar eH/mc$.

The spectrum of the 1 November 1979 burst is shown in Fig. 1b, c. In contrast to the 7 March 1979 event, the absorption line is observed here only at the initial burst stage (Fig. 1b). At the later stage (Fig. 1c) the line disappears—an evolution typical of many other bursts which may be associated with the heating of the absorption region during the burst development. The increase in emission at the cyclotron frequency in this region due to heating can cancel absorption in the spectrum of the γ -burst. In the 1 November 1979 burst the spectral distribution in the continuum also undergoes evolution. The spectrum becomes harder and the temperature kT increases from 240 to 480 keV.

An interesting feature of the absorption line in the spectrum of the 1 November 1979 event is an intensive wing at lower energies which causes a marked broadening of the equivalent line width. Apparently, this can be attributed to a spatial variation in magnetic field in the source. Such an interpretation also holds for many other bursts, for example, those with the spectra shown in Fig. 1d–f. Here the absorption line spreads into a band because of a substantial change in magnetic field within the source.

Figure 2a, b shows spectra with absorption lines obtained for the bursts on 29 March 1979 and 24 May 1979. Note that these bursts were recorded by the Venera 11 detector whose energy

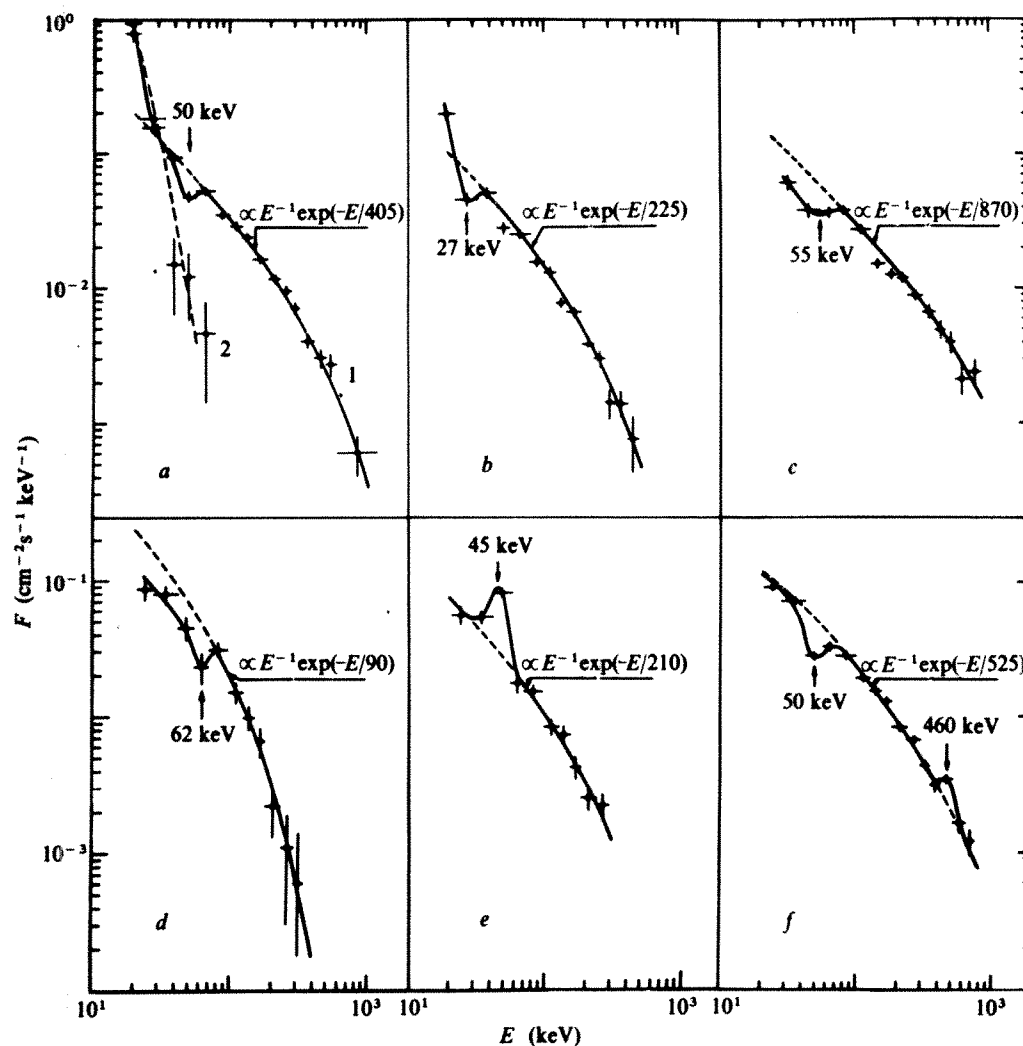
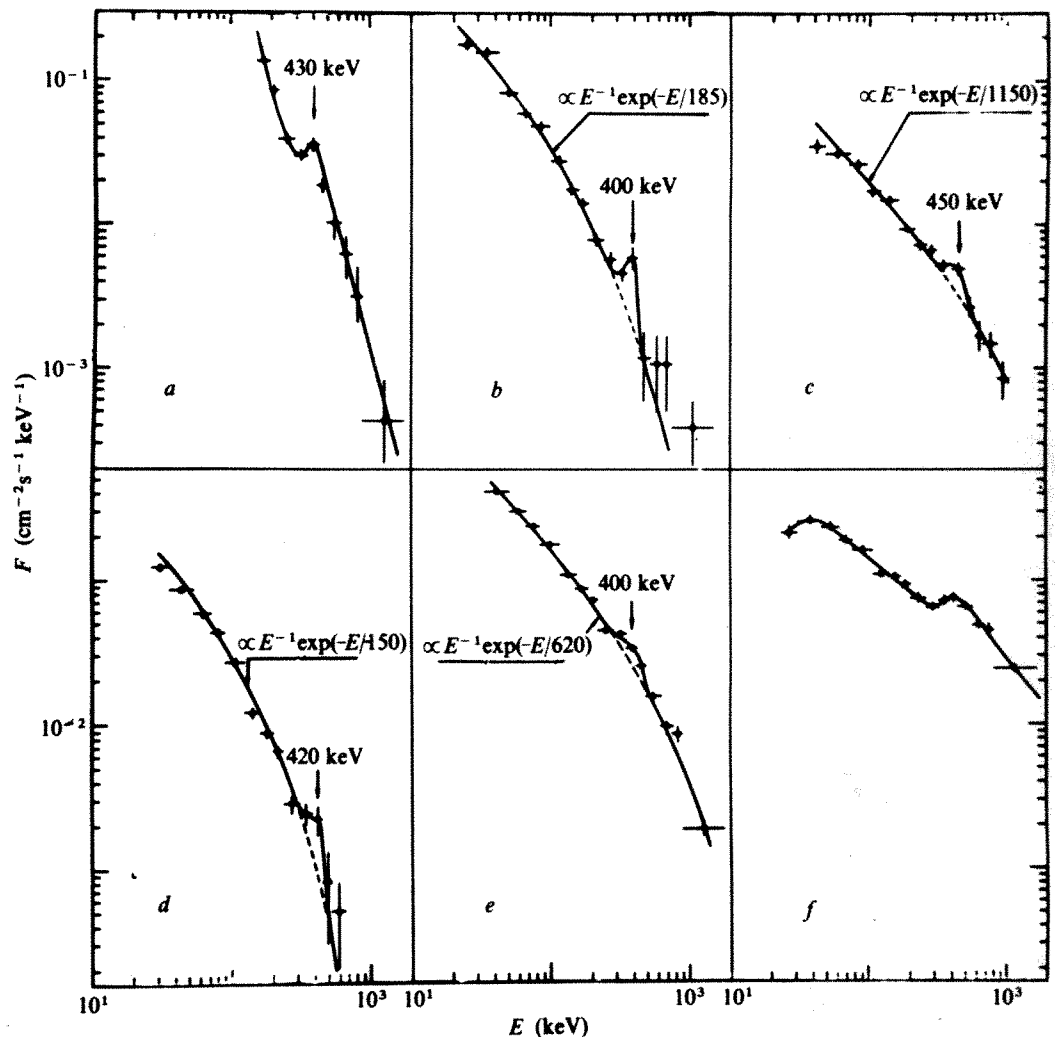


Fig. 2 The energy spectra of γ -bursts. a, 29 March 1979 (1, the spectrum of the initial stage of the burst; 2, the spectrum obtained for the deep minimum between the burst peaks); b, 24 May 1979; c, 31 July 1979 (initial stage); d, 6 October 1979 (initial stage); e, 26 May 1979; f, 22 June 1979.

Fig. 3 The energy spectra of γ -bursts with emission lines. *a*, 5 March 1979 (the spectrum of the short initial pulse); *b*, 18 September 1978; *c*, 18 April 1979; *d*, 16 January 1979; *e*, 4 November 1978 (initial stage); *f*, 19 November 1978 (initial stage).



threshold lies at ~ 17 keV in contrast to the other five detectors with a threshold of ~ 30 keV. This enabled us to obtain additional information. The radiation level in the first, the softest channel of the spectra of these bursts, is higher than normal and remains nearly constant during the burst. The 29 March 1979 event consists of many peaks and lasts about 40 s (ref. 14). The last measured spectrum (2 in Fig. 2a) falls on the deep minimum between the peaks—its shape corresponds to radiation from a substantially colder plasma with temperature $kT \approx 8$ keV. A comparison of the two spectra in Fig. 2a shows that the soft component of the radiation is present during the entire burst. This suggests that the bursts on 29 March 1979 and 24 May 1979 come from two-component sources. It is unclear whether the appearance of the absorption lines can be associated with the presence in the source of a substantially more stationary and colder plasma.

Figure 2c, d shows absorption at the initial stages of the 31 July 1979 and 6 October 1979 bursts which also undergoes evolution, disappearing at later stages. All these graphs demonstrate the similarity between the patterns of absorption in the spectra with substantially different spectral distributions in the continuum.

The burst on 26 May 1979 (Fig. 2e) revealed an emission line at 45 keV. From our data, this spectrum is the only reliable evidence of emission at the cyclotron frequency ω_H . Figure 2f shows a spectrum of the 22 June 1979 burst which is remarkable in that it contains in addition to the cyclotron absorption line and an emission line at 460 keV which is apparently due to the annihilation radiation. The energy of this line was reduced $(1 - 2GM/c^2R)^{1/2}$ times as a result of gravitational redshift in the field of a neutron star.

The same emission lines were also observed in several other γ -ray bursts (Fig. 3). These data show that the relative intensity

of the emission lines, just as that of the absorption lines, can vary in a time shorter than the burst duration. The line emission seems most likely to appear at the burst stages characterized by the highest temperature of radiation in the continuum. Indeed, the line at ~ 430 keV in the burst on 5 March 1979 can be seen only in the hard spectrum of the short initial pulse¹¹. A broad line at ~ 400 keV simultaneous with hard radiation in the continuum can only be observed at the initial stage of the 4 November 1978 burst. The event on 19 November 1978 is distinctive because of its extremely strong evolution of the spectral distribution⁹. The very hard spectrum with a broad emission feature at 300–800 keV (Fig. 3f) is observed during the first 4 s of the burst. The resolution in this region of two lines at 420 and 740 keV (ref. 12) apparently represents the only case of observation in a γ -ray burst of a line above 500 keV. The temporal evolution of the emission lines in the 18 September 1978, 18 April 1979, and 16 January 1979 bursts is less pronounced.

Tables 1 and 2 contain data on lines in the γ -ray bursts, showing the date of the event, the time of the beginning of recording T_0 , and the energy flux S (>30 keV) integrated over the burst duration. For the emission lines the mean energy $\bar{E}$, the mean photon flux F , the energy flux S_e emitted in the line, and the ratio S_e/S are given. The parameters for the absorption lines are the mean energy $\bar{E}$, the relative line depth h in units of the continuum intensity, the estimated equivalent line width $W = \int h dE$, the estimated fraction of absorbed energy flux in the burst S_a , and the ratio S_a/S . If evolution of a line was observed, the values of S and S_e/S , or S_a/S calculated for the burst stage with the line are also given in brackets.

The figures and tables show that the lines in different bursts have fairly similar characteristics. Emission lines appear predominantly in the interval 400–460 keV and the energy

contained in these constitute 5–10% of the energy of the burst. Absorption lines are observed in the 30–70-keV region and the equivalent line widths constitute 5–15 keV, thus 1–2% of the energy of the burst is absorbed. In some cases the lines become smeared strongly and the integral absorption in the soft region of the spectrum increases. The temporal evolution of lines in different bursts also reveals common features. Such a similarity supports our interpretation of the spectral features in the γ -burst spectra as cyclotron absorption lines and redshifted annihilation lines. Thus an overall analysis of our data leads to a conclusion that γ -bursts are produced in regions with a gravitational potential of $\sim(0.1\text{--}0.2)c^2$ and magnetic field $\sim(2\text{--}6) \times 10^{12}$ Oe near the surface of neutron stars. Strong support for these neutron stars being actual components of binaries¹⁶ is the discovery of a similar cyclotron absorption line at 20 keV in the spectrum of the pulsed hard X-ray flux¹⁷ from 4U0115+63.

Received 28 October; accepted 21 December 1980.

1. Ramaty, R. & Lingenfelter, R. E. *Nature* **278**, 127–132 (1979).
2. Vedrenne, G. *Adv. Space Explor.* **7**, 85–99 (1979).
3. Gnedin, Yu. N. & Sunyaev, R. A. *Astron. Astrophys.* **36**, 379–394 (1974).
4. Canuto, V. & Chiu, H. Y. *Space Sci. Rev.* **12**, 3–74 (1971).
5. Trümper, J. *et al. Astrophys. J. Lett.* **219**, L105–L110 (1978).
6. Ling, J. C. *et al. Astrophys. J.* **231**, 896–905 (1979).
7. Jacobson, A. S. *et al. Gamma Ray Spectroscopy in Astrophysics* (eds Cline, T. L. & Ramaty, R.) 228–251 (NASA TM 79619, 1978).
8. Leventhal, M., MacCallum, J. C. & Watts, A. C. *Astrophys. J.* **216**, 491–502 (1977).
9. Mazets, E. P. *et al. Kosm. Issled.* **17**, 812–819 (1979).
10. Mazets, E. P. *et al. Preprint no. 599* (A. F. Ioffe Physico-Technical Institute, 1979).
11. Mazets, E. P. *et al. Pisma Astr. Zh.* **5**, 307–312 (1979); *Nature* **282**, 587–589 (1979).
12. Teegarden, B. J. & Cline, T. L. *Astrophys. J. Lett.* **236**, L67–L70 (1980).
13. Mazets, E. P. *et al. Pisma Astr. Zh.* **6**, 706–711 (1980).
14. Mazets, E. P. *et al. Preprints nos 618 & 637* (A. F. Ioffe Physico-Technical Institute 1979).
15. Mazets, E. P. *et al. Preprint no. 662* (A. F. Ioffe Physico-Technical Institute, 1980); *Astrophys. Space Sci.* (in the press).
16. Mazets, E. P. & Golenetskii, S. V. *Preprint No. 632* (A. F. Ioffe Physico-Technical Institute, 1979); *Astrophys. Space Sci.* (in the press).
17. Wheaton, Wm. A. *et al. Nature* **282**, 240–243 (1979).

Observations of a circumstellar shell around the OH/IR star OH127.8–0.0

R. S. Booth, A. J. Kus*, R. P. Norris & N. D. Porter

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

The 1,612 MHz OH maser emission associated with cool long-period variable stars is generally characterized by two main regions of emission separated in velocity by $\sim 10\text{--}50$ km s^{−1}. This has been taken to indicate that the OH lies in an expanding shell surrounding the star, the main peaks originating at the near or far sides of the shell along the line of sight where the maser column lengths may be greatest. Here we present the first high-angular resolution measurements of the OH emission from such a source which demonstrates dramatically that this picture is correct.

The 1,612-MHz OH maser source, OH127.8–0.0 (hereinafter called OH127.8) was discovered in a survey by Bowers¹. Like many other sources found in such surveys^{2,3}, it shows the OH characteristics of the long period variable stars but has no optical identification although it is probably coincident with an IR star⁴. These objects have been designated Type II OH/IR stars and are assumed to be long period variable stars which are either at too large a distance for optical identifications or have circumstellar shells which are too thick for the visible radiation to penetrate (their properties are discussed elsewhere^{5,6}).

Observations of the OH maser radiation from OH127.8 were made between late 1978 and early 1980 with three single baseline interferometers of the Jodrell Bank telescope network, before the completion of the new Multi Telescope Radio Linked

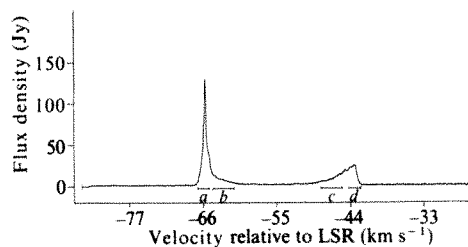


Fig. 1 The 1,612 MHz OH spectrum of OH127.8. The bracketed velocity intervals *a*, *b*, *c* and *d* represent those ranges of velocity for which the spatial distribution of the maser emission has been determined (see Fig. 2).

Interferometer (MTRLI)⁷ and as a prelude to its use as a line instrument.

These observations of OH maser radiation form part of a larger programme of observations of several such objects using the MkIA–Knockin, the MkIA–Defford and the MkIA–MkIII interferometers respectively⁸. On each occasion the usual radio link technique was used and the data were cross-correlated in the 1,024-channel digital spectrometer to give a 512 point cross-correlation spectrum in the manner described by Norris and Booth¹⁸.

The autocorrelation spectrum of OH127.8 is shown in Fig. 1. The special feature of this source is that the narrow spectral component at -65.7 km s^{−1} was virtually unresolved on all baselines and could therefore be used as phase reference not only to correct the phase on all the other channels of the cross-correlation spectrum but also to combine the three sets of interferometer data. After combination and normalization the new data set was split into velocity intervals and a series of maps produced, each corresponding to a particular range of velocities in the masing gas. The procedure, described fully elsewhere⁹ consisted essentially of Fourier inversion and cleaning⁹.

Four of the maps are reproduced in Fig. 2*a–d*—they represent the spatial distribution of the masing gas in the velocity intervals *a*, *b*, *c* and *d* indicated in Fig. 1. Figure 2*a* is a map of the distribution of the OH gas with velocities close to the reference feature and in the range -65 to -66.5 km s^{−1} (region *a* in Fig. 1). It shows that the material within this velocity interval is concentrated in a small region of angular extent ~ 0.4 arc s and is

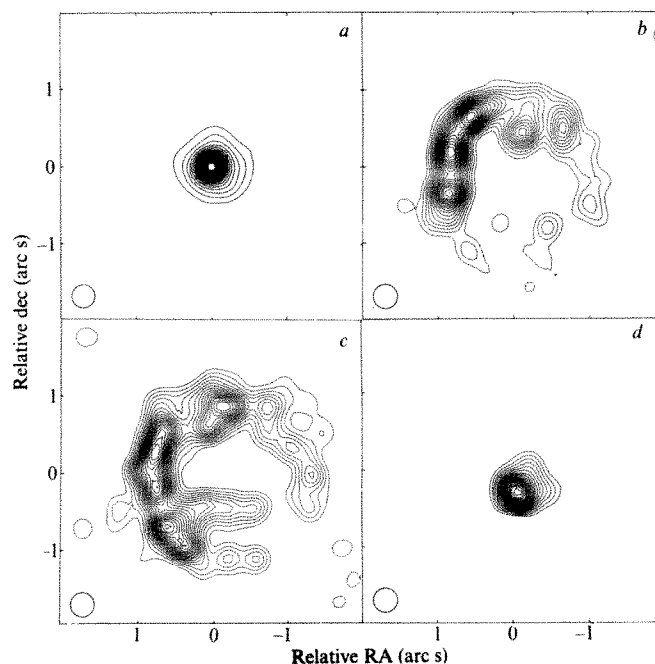


Fig. 2 *a–d*, Maps of the spatial distribution of the integrated OH emission in the velocity intervals marked *a* to *d* respectively in Fig. 1. In each map the contour interval is 5% of the peak integrated emission in that velocity interval and the lowest contour is $\sim 10\%$.

* Permanent address: Nicolaus Copernicus University, Gagarina 11, 87-100 Toruń, Poland.

dominated by the -65.7 km s^{-1} component. Figure 2d shows the distribution of OH in the corresponding redshifted velocity interval (-43 to -43.7 km s^{-1}) and is clearly similar in extent although slightly offset in position.

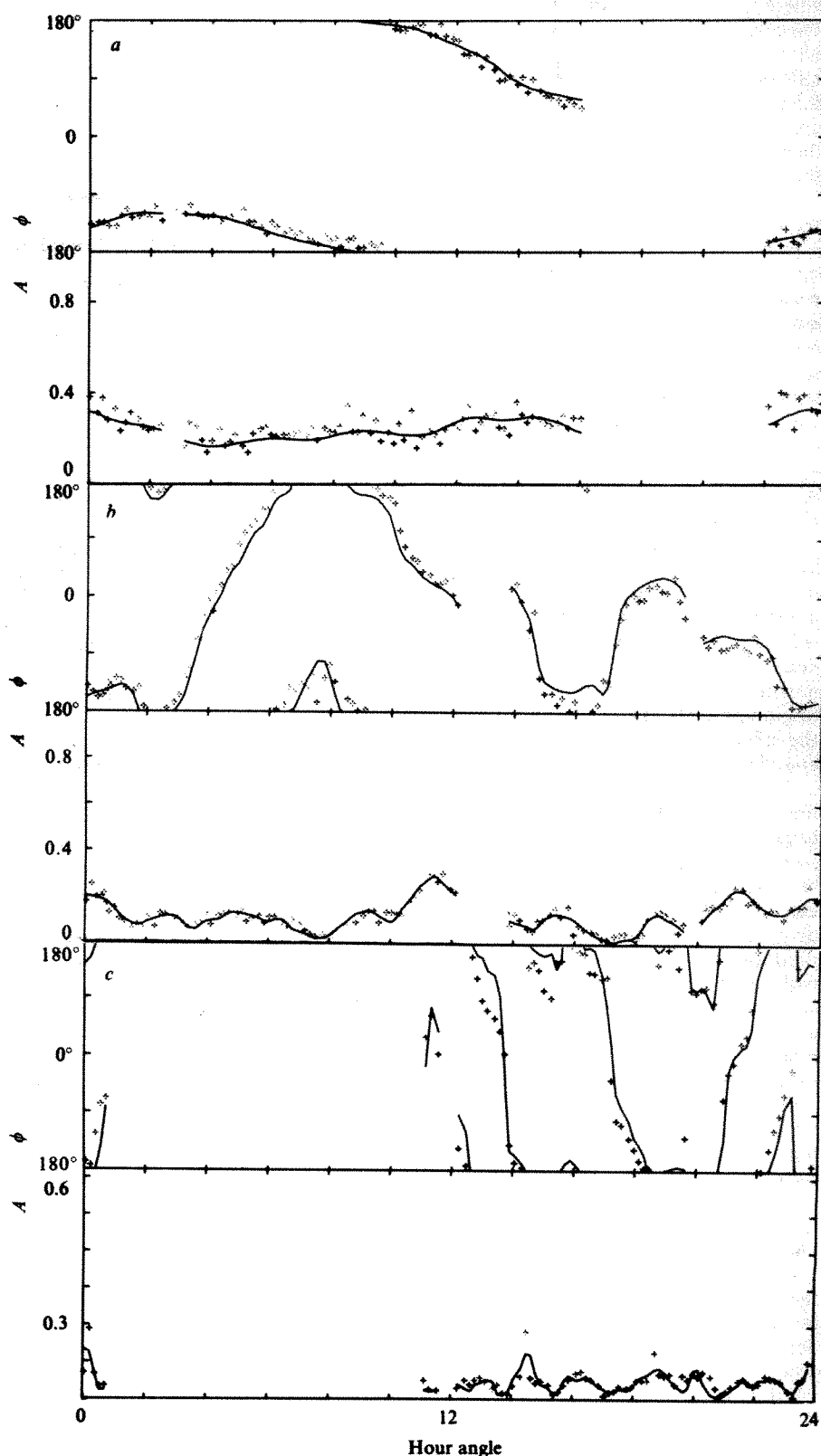
Figure 2b, c shows maps of the weaker OH emission in velocity regions closer to the stellar velocity (-62.2 to -64.7 km s^{-1} and -45.2 to -47.5 km s^{-1} respectively). Both have a similar annular form with angular diameter $\sim 2 \text{ arc s}$. The map centres are the same in each case and correspond to the position of the reference feature.

The distributions found are exactly those expected of a uniformly expanding circumstellar shell¹⁰ of OH. The strongest

maser components have small angular diameters and are situated close to the line of sight through the centre of the shell along which the gain paths are largest, and the weaker emission delineates those portions of the shell where constant velocity path lengths along the line of sight will be shorter, giving lower maser gain. At the edges of the shell the velocity gradient is such that the line of sight emission is reduced to zero.

Note that the spectrum of the source varied slightly in intensity between observations thus we could not be sure of the relative amplitudes to better than 10%. However, as the fits to the data were good, we believe that our normalization was correct; as an example, in Fig. 3 we show a plot of the data and

Fig. 3 A plot of the visibility phase (ϕ) and amplitude (A) of the OH emission in the velocity range b of Fig. 1 together with the same data (solid line) derived from Fourier inversion of the map plotted in Fig. 2b. Baselines, *a*, MkIA-MkII; *b*, MkIA-Knockin; *c*, MkIA-Deffprd.



corresponding points produced by Fourier inverting the map of Fig. 2b. As an extra check on this effect, we varied the amplitudes by 10% on each baseline and repeated the fitting procedures. In all cases, although the detailed distribution changed, the overall shell pattern remained, it being strongly dependent on the phase data which were always referenced to the unresolved point component at -65.7 km s^{-1} .

Assuming the uniformly expanding shell model¹⁰, the radius at which the line of sight component of the expansion velocity is v , $a(v)$ is given by

$$a(v) = R[1 - (v/v_e)^2]^{1/2}$$

where v_e is the expansion velocity and R the total shell radius. In the case of OH127.8, the expansion velocity (half the interval between OH peaks in Fig. 1) is $\sim 11 \text{ km s}^{-1}$. The rings of emission in Fig. 2b, c have mean velocity relative to the star of 8.5 km s^{-1} and angular diameter ~ 2 arcs. Thus the total shell angular diameter is ~ 3.1 arcs which corresponds to a linear diameter of $1.5 \times 10^{17} \text{ cm}$ if we take the kinematic distance to the star to be 3.3 kpc^{11} .

A technique used recently to determine the linear diameter of the OH shells surrounding long period variable stars¹² involves careful observation of the OH variability when a phase difference can be observed in the light curve of the blueshifted emission from the nearside of the shell, relative to the redshifted emission from the far side. Assuming that the central star is the source of pump radiation, the OH emission from the entire shell should vary in unison and the observed phase lag represents the shell crossing time. Phase lag measurements by Jewell *et al.*¹³ and Herman and Habing¹⁴ show that diameters, d , lie in the range $10^{16} \text{ cm} < d < 10^{18} \text{ cm}$. Although these groups have not sufficiently monitored OH127.8 to obtain a definite phase lag result on this source, our value of diameter lies well within the range of measured values.

The mass of the circumstellar shell and particularly the mass loss rate from long period variable stars are of great importance in understanding stellar evolution and the chemical composition of the interstellar gas. Observations of the circumstellar masers will provide information on the masses involved if the gas density in the shell can be determined. This involves a determination of the OH column density through consideration of the brightness temperature of the maser, its degree of saturation and the efficiency of its pump mechanism, and an estimate of the abundance of OH relative to H_2 . In the case of OH127.8, individual maser features are unresolved with our longest baseline and they may be as small as 30 m arc s (ref. 11). Thus their brightness temperatures are $\sim 10^9 \text{ K}$. An efficient $35\text{-}\mu\text{m}$ IR pumping scheme for the masers in circumstellar envelopes has been developed by Elizur *et al.*¹⁵ and their calculations show that the $1,612 \text{ MHz}$ transition is inverted and saturated for OH densities $n_{\text{OH}} \geq 1 \text{ cm}^{-3}$. Brightness temperatures of $\sim 10^9$ are produced for OH column densities $\sim 10^{17} \text{ cm}^{-2}$, which agrees with our observations. A detailed model of the circumstellar OH abundance has been developed by Goldreich and Scoville¹⁶ who argue that $n_{\text{OH}} \geq 1 \text{ cm}^{-3}$, at a radius of $\sim 10^{16} \text{ cm}$ when the molecular hydrogen number density is $\sim 2 \times 10^4 \text{ cm}^{-3}$. This value, together with our measured linear diameter, gives a value of $\sim 10^{-1} M_{\odot}$ for the mass of the shell and a mass loss rate $\sim 8 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$. Although such a mass loss rate is extremely high, it agrees well with rates derived for a sample of OH/IR stars from measurements of their IR properties, luminosity and dust shell opacities together with the outflow velocities¹⁷. Clearly, stars losing mass at these rates must be short-lived phenomena.

Finally OH/IR stars generally have angular sizes which lie in the resolution range of the MTRLI and observations are already being made using this instrument. As well as measuring the detailed structure and mass loss rates of these fascinating objects, a combination of the angular size of their shells from interferometry and the direct (phase lag) measurement of their diameter will lead to valuable independent estimates of their distances.

We acknowledge the assistance of Dr B. Anderson with some of the equipment and Drs P. N. Wilkinson, R. G. Noble and T. J. Cornwell for their advice on the mapping procedures. A.J.K. received a British Council support grant.

Received 17 December 1980; accepted 9 February 1981.

1. Bowers, P. F. *Astr. Astrophys. Suppl.* **31**, 127–145 (1978).
2. Johanson, L. E. B. *et al. Astr. Astrophys.* **54**, 323 (1977).
3. Baud, B. *et al. Astr. Astrophys.* (in the press).
4. Engels, D. *Astr. Astrophys. Suppl.* **36**, 337 (1979).
5. Snyder, L. E. in *Interstellar Molecules* (ed. Andrew, B. H.) (Reidel, Dordrecht, 1980).
6. Elitzur, M. in *Proc. Workshop on phys. Processes in Red Giants, Erice 1980* (eds Renzini, A. & Iben, I.) (in the press).
7. Davies, J. G., Anderson, B. & Morison, I. *Nature* **288**, 64 (1980).
8. Porter, N. D. *et al.* (in preparation).
9. Hogbom, J. A. *Astr. Astrophys. Suppl.* **15**, 417 (1974).
10. Reid, M. J. *et al. Astrophys. J.* **214**, 60 (1977).
11. Bowers, P. F. *et al. Astrophys. J.* **242**, 1088 (1980).
12. Schultz, G. V., Sherwood, W. A. & Winnberg, A. *Astr. Astrophys.* **63**, L5 (1978).
13. Jewell, P. R., Webber, J. C. & Snyder, L. E. *Astrophys. J.* **242**, L29 (1980).
14. Herman, J. & Habing, H. in *Proc. Workshop on phys. Processes in Red Giants, Erice 1980* (eds Renzini, A. & Iben, I.) (in the press).
15. Elitzur, M., Goldreich, P. & Scoville, N. *Astrophys. J.* **205**, 384 (1976).
16. Goldreich, P. & Scoville, N. *Astrophys. J.* **205**, 144 (1976).
17. Werner, M. W. *et al. Astrophys. J.* **239**, 540 (1980).
18. Norris, R. P. & Booth, R. S. *Mon. Not. R. astr. Soc.* (in the press).

Fragmentation of a meteor in near-Earth space

M. A. Hapgood & P. Rothwell

Department of Physics, The University,
Southampton SO9 5NH, UK

Fragments of a Perseid meteor, which broke up well above the Earth's atmosphere, probably in collision with a dust particle in near Earth space, have been observed with a two-station low-light level television camera system. Three Perseid meteoroids, with apparent magnitudes between 2.0 and 3.5, arrived on near parallel trajectories within 1.3 s of each other on the morning of 12 August 1977 (03.05.21–03.05.23 UT). The probability of these meteoroids arriving by chance so close together in time and direction is very small ($\sim 3 \times 10^{-11}$). Assuming the three meteoroids were fragments of a larger body, we show here that the parent meteor must have broken up at least 1,700 km above the Earth's surface, well above the altitude at which meteoroids normally start to burn up in the Earth's upper atmosphere (100–150 km). An estimate of the size of the parent meteor has been made from the observed magnitudes of these fragments. Assuming the parent meteoroid was spinning with an angular velocity between 100 and 1,000 rad s^{-1} , an upper limit to the distance of breakup point from the Earth's surface can be placed at between 7×10^5 and $7 \times 10^6 \text{ km}$ (0.004–0.04 AU). We show here that the probabilities of collision and breakup of the parent Perseid with three classes of dust particle in near Earth space—interplanetary, magnetospheric and lunar ejecta—at the time of our observation in 1977, were $\sim 3 \times 10^{-6}$, 3×10^{-6} and 5×10^{-7} respectively. We suggest therefore that if the Perseid parent comet (Swift–Tuttle) passes close to the Earth when it returns to perihelion in the next year or two, giving rise to a spectacular display of the Perseids, we should then be able to see many of these associated meteor events, and to study meteor collision with dust particles in near Earth space.

Table 1 Meteor parameters

	Meteor 1	Meteor 2	Meteor 3
Geocentric radiant:			
RA (α)	3 h 26 min	3 h 31 min	3 h 19 min
Dec (δ)	+53.7°	+54.2°	+53.8°
Geocentric velocity (km s^{-1})	59 $\pm$ 2	59 $\pm$ 2	56 $\pm$ 3
Apparent magnitude	2.0 $\pm$ 0.5	2.1 $\pm$ 0.5	3.5 $\pm$ 0.5
Magnitude at zenith	0.5	0.6	1.8
Mass (10^{-5} kg)	13	12	4

The meteor observations were made with low-light level TV cameras and associated video recording equipment at two sites in the south of England; Southampton (50°50' N, 1°24' W) and Harwell (51°35' N, 1°20' W). An image isocon camera, with a 30° × 23° field of view, was used at Southampton and a silicon intensified target (SIT) camera, with a 20° × 15° field of view, was used at Harwell. Both cameras looked west at a common volume centred above South Wales at a height of ~100 km. Fifty pictures per second were recorded on video tape; these pictures could be played back slowly frame by frame. Typically a meteor appeared in ~10 successive frames.

The positions in space of the three associated Perseids were determined by triangulation against the fixed star background using observations from the two sites. Figure 1 shows the variation of the altitude of the lower tip of the meteor trails with time; Fig. 2 shows the positions of the three trails projected onto the ground. The velocities and radiant of the three meteors are listed in Table 1; they all lie close to the mean Perseid values¹ (geocentric velocity 59.4 km s⁻¹, geocentric radiant $\alpha = 3$ h 05 min, $\delta = 57.4^\circ$). The error in the determination of each velocity is ~2 km s⁻¹; the error in the position of one radiant relative to another is ~1°. The apparent magnitude of each meteor was determined by comparing its brightness with the brightness of stars in the field of view. The meteor magnitudes, corrected to the zenith², are also listed in Table 1, together with the masses, estimated from the zenith magnitudes, following Hughes³.

The three meteors arrived within a time interval of 1.3 s and on trajectories which were parallel to within 1°, that is, within a solid angle of ~10⁻³ sr. In contrast the average meteor rate in the field of view of our cameras during the Perseid meteor shower was ~6 h⁻¹; the radiant of these meteors were spread over a wide region of the sky ~12° across (solid angle ~0.15 sr). Thus the probability of observing by chance three meteors arriving so close together in time and direction was ~3 × 10⁻¹¹.

By combining the masses of the three meteors (see Table 1) we can estimate that the mass of the parent body was ~3 × 10⁻⁴ kg. If we assume that the parent body was roughly spherical and had a typical Perseid density⁴ of 290 kg m⁻³ its radius was ~6 mm.

We can estimate the distance s from the Earth back to the breakup point since $s = v^2 \Delta t / \Delta v$, where v is the mean velocity of the fragments, Δv is their velocity dispersion and Δt is the time interval between the arrival of the fragments. The mean velocity ($v = 58.5 \pm 1.3$ km s⁻¹) and the time interval ($\Delta t = 1.30 \pm 0.02$ s, see Fig. 1) are known quite accurately. However, the velocity dispersion Δv lies within the errors of measurement so the error in the mean velocity sets an upper limit on Δv . We take $\Delta v_{\max}$ to be 2.6 km s⁻¹, which gives a value of the minimum distance to breakup point, $s_{\min}$ of 1,710 km. We can make a second and independent estimate of $s_{\min}$ from the observation that the

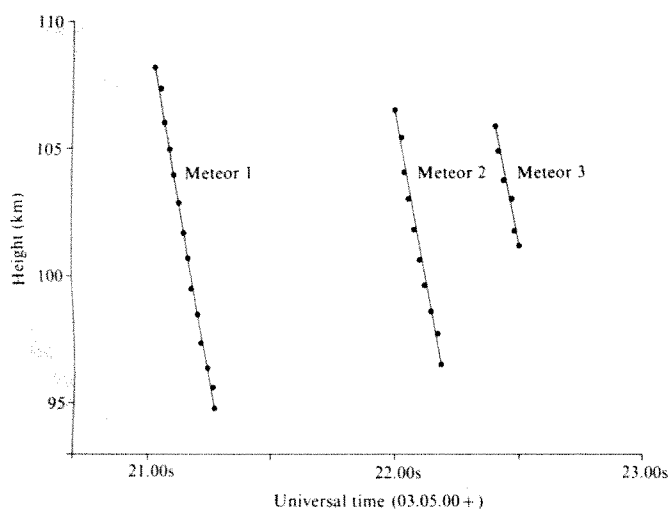


Fig. 1 The variation of meteor heights as a function of time.

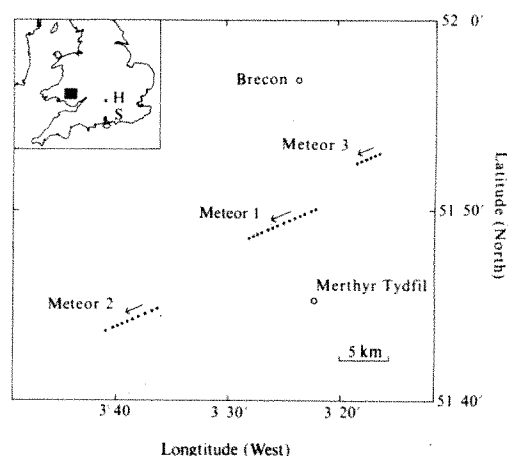


Fig. 2 The horizontal motion of the meteors. The position of each meteor is shown at time intervals of 20 ms (1 TV picture). The inset shows the area (shaded) of the main map and the positions of the two observing sites: Southampton (S) and Harwell (H).

meteor fragments travelled on paths which were parallel to within the errors of measurement (1°) and were ~30 km apart at 105-km altitude. Taking the maximum angle of dispersion between the trajectories to be 1° we obtain $s_{\min} = 1,720$ km which agrees well with the value deduced from the velocity dispersion. This puts the fragmentation point outside the Earth's atmosphere, thus a high velocity collision between the parent body and a small dust particle would be the likely cause of fragmentation.

We can set an upper limit to the distance of the breakup point from the Earth if we can set a minimum value for the velocity dispersion of the fragments ($\Delta v_{\min}$). In general we would expect a significant velocity dispersion to be generated by the collision. However, in the case of a collision with a small dust particle which is only just large enough to cause fragmentation, almost all the kinetic energy of the collision would be dissipated in breaking up the parent meteoroid. In such cases the only contribution to the velocity dispersion is that arising from the rotation of the parent body. Thus we may put $\Delta v_{\min} \sim \omega r$ where ω is the angular velocity of the parent body and r is its radius. Theoretical studies indicate that meteoroid spin may be produced by collisions⁵ and by solar radiation effects⁶. These theories indicate that angular velocities of 100–1,000 rad s⁻¹ may be achieved within the lifetime of the Perseid stream (~3,000 yr (ref. 7)). This implies a value for $\Delta v_{\min}$ in the range 0.6–6 m s⁻¹ giving a value $s_{\max}$ in the range 7–70 × 10⁵ km (0.004–0.04 AU). Thus we may conclude that the parent body broke up in space close to the Earth.

The minimum mass of dust particle needed to fragment the parent body can be estimated from the results of the hyper-velocity impact experiments of Moore and Gault⁸. They found that a body of mass M would be fragmented by the impact of a particle of mass m and relative velocity v km s⁻¹ when $m > M/250 v^2$. If we apply this formula to the present case we find that a dust particle of minimum mass 3×10^{-10} kg would be required to give fragmentation. However, note that the hyper-velocity impact experiments from which the formula was derived were carried out on basalt targets with relative velocities up to 7 km s⁻¹. The use of the Moore and Gault⁸ formula involves considerable extrapolation in velocity and assumes that the meteoroid material fragments in similar conditions to basalt.

We may estimate the probability of observing a group of associated meteors by calculating the probability of the parent body suffering a collision. Dohnanyi⁹ has calculated the collisional lifetimes T_c of meteoroids in several meteor streams, with interplanetary dust particles, taking into account both the meteoroid cross-section and the number density of dust particles which are, according to Moore and Gault⁸, large enough to cause fragmentation. He finds that, for the Perseids, $T_c = 1.8 \times 10^4 M^{1/3}$ yr where M is the meteoroid mass in kg. Thus

with $M = 3 \times 10^{-4}$ kg, $T_c = 1,200$ yr (4×10^{10} s). The probability of collision with an interplanetary dust particle within 7×10^6 km of the Earth, for a meteor of mass 3×10^{-4} kg travelling at 58.5 km s $^{-1}$, is $\sim 3 \times 10^{-6}$. This is a small probability but still $\sim 10^5$ times higher than the chance probability of three meteors arriving so close together in time and direction.

Fechtig *et al.*¹⁰ have reported that the flux of dust particles in the mass range 10^{-12} – 10^{-18} kg observed from the HEOS 2 satellite within the Earth's magnetosphere, is ~ 100 times greater than the flux in interplanetary space. If a similar enhancement exists for masses of $\sim 10^{-10}$ kg, the collisional lifetime for a Perseid in this region would be reduced to $\sim 4 \times 10^8$ s. It takes $\sim 1,000$ s for a Perseid to cross the magnetosphere, thus the probability of a catastrophic collision with a magnetospheric dust particle is therefore $\sim 3 \times 10^{-6}$, which is comparable with the probability of a collision with a dust particle in interplanetary space.

Fechtig *et al.*¹⁰ also observed groups of particles from the HEOS 2 satellite which are believed to be high velocity ejecta from the impact of large meteoroids on the surface of the Moon. The flux of these particles was about one-tenth of the flux of dust particles in interplanetary space. During meteor showers the number of lunar ejecta increases. Alexander and Corbin¹¹ have calculated that the mass flux ejected from the lunar surface by the Perseid meteor shower is ~ 30 times that ejected by sporadic meteors. They have also shown that when the Moon is in the last quarter, a large percentage of the lunar ejecta are directed towards the Earth. At the time of our observation of 12 August 1977, the lunar phase angle from full Moon was 146° (in the last quarter). The flux of lunar ejecta near the Earth may, therefore, have been about three times the interplanetary dust flux. If these ejecta particles are concentrated in a region comparable in size to the Moon's orbit (4×10^5 km) then the probability of a catastrophic collision with a lunar ejecta particle is $\sim 5 \times 10^{-7}$.

The three Perseid meteors we observed ablating in the Earth's atmosphere, closely associated in space and time, are likely to be fragments of a larger Perseid meteoroid of mass 3×10^{-4} kg which broke up between $1,700$ and 7×10^6 km from the Earth's surface. The probability P_1 that they are fragments from a single meteor colliding with a dust particle in interplanetary space is $\sim 10^5$ times greater than the probability of a random association between the three meteors. The probability P_2 that the parent Perseid collided with a dust particle in the higher density dust environment of the Earth's magnetosphere is comparable with P_1 . The probability P_3 that the parent Perseid collided with a moon dust particle ejected when the Moon was bombarded with meteors from the Perseid shower, is $\sim$ one-sixth of P_1 .

If the Perseid broke up on collision with a dust particle we were lucky to observe this event. The chance of an individual meteor breaking up close to the Earth was of the order of a few in a million and we recorded less than 100 meteors during the 1977 Perseid meteor shower. During spectacular displays of meteors, when the parent comet passes close to the Earth, the meteor rate can rise by two or three orders of magnitude. The dust population in the magnetosphere and the density of lunar ejecta in near Earth space should rise and the chance of an individual meteoroid breaking up close to Earth on collision with a dust particle could increase up to a few in a thousand. In this case several events of the type we have described could be observed during the meteor display. The parent comet of the Perseid meteor stream (Swift–Tuttle, 1862 III) is expected to return to perihelion sometime in the next year or two. There has been speculation that the passage of the comet may give rise to a spectacular Perseid display. We may then be able to make further studies on the breakup of Perseid meteors in near Earth space.

We thank Dr O. Royrvik for observational assistance and Dr J. V. Jelley for use of his observing site at Harwell.

Received 28 August 1980; accepted 16 February 1981.

1. Cook, A. F. *NASA Contractor Rep.* No. 2109, 153–166 (1972).
2. Lovell, A. C. B. *Meteor Astronomy*, 131 (Oxford University Press, 1954).

3. Hughes, D. W. in *Cosmic Dust* (ed. McDonnell, J. A. M.) 151 (Wiley, Chichester, 1978).
4. Hughes, D. W. in *Cosmic Dust* (ed. McDonnell, J. A. M.) 164 (Wiley, Chichester, 1978).
5. Hawkes, R. L. & Jones, J. *Mon. Not. Roy. astr. Soc.* **185**, 727–734 (1978).
6. Paddack, S. J. & Rhee, J. W. *Geophys. Res. Lett.* **2**, 365–367 (1975).
7. Southworth, R. B. *Smithsonian Contr. Astrophys.* **7**, 299–303 (1963).
8. Moore, H. J. & Gault, D. E. *U.S. geol. Survey A. Progr. Rep.* 127–150 (1965).
9. Dohnanyi, J. S. *J. geophys. Res.* **75**, 3468–3493 (1970).
10. Fechtig, H., Grün, E. & Morfill, G. *Planet. Space Sci.* **27**, 511–531 (1979).
11. Alexander, W. M. & Corbin, J. D. *Submicron Lunar Ejecta in the Magnetosphere Associated with Meteor Showers* (COSPAR, Budapest, 1980).

Atomic motion on the surface of a cadmium telluride single crystal

R. Sinclair, T. Yamashita & F. A. Ponce

Department of Materials Science and Engineering, Stanford University, Stanford, California 94305, USA

Cadmium telluride (CdTe) is an interesting semiconductor compound whose photoelectronic properties make it suitable for solar cell applications. We report here a study of defects in CdTe, carried out by lattice resolution imaging in a transmission electron microscope (TEM), during which observations were made of the motion of image spots at the edge of a specimen. As the spots are related to the projected atomic positions in the structure, this behaviour indicates atomic motion at the CdTe surface. Individual spot jumps occurred several times per second and one could watch the formation and decay of small crystalline facets. The possibility thus arises for surface studies, at the atomic level, in conjunction with direct crystal lattice imaging, allowing investigation of the interaction between surface atoms and their substrate.

CdTe has the zinc blende structure which is best imaged by high-resolution TEM with a $\langle 110 \rangle$ direction parallel to the electron beam. The projection of the structure in this orientation is shown in Fig. 1a. Adjacent columns of cadmium and tellurium atoms are separated by 1.6 Å, but their resolution is only achieved in very restrictive experimental circumstances^{1,2}. More commonly, bright spots are seen in the image (for example, Fig. 1b), the distribution of which corresponds to the arrangement of the projected lattice points. Such images are consistent with comparable images of elemental germanium and silicon which have the related diamond cubic structure^{3–5}. Theoretical image calculations, discussed later, establish the relationship between image spots and the crystal structure.

Microscopy was performed using a Philips EM400 TEM equipped with a high-magnification objective lens ($C_o = 1.1$ mm at 120 kV) and a LaB₆ electron source. This lens has no tilting capability. Therefore the single crystals were sectioned as closely as possible to the desired orientation. Thin foils were obtained by standard ion-milling procedures. A region of the specimen close to the perfect 110 zone axis orientation was found by translating the specimen in the microscope. The present images were taken by allowing 36 beams through the objective aperture (Fig. 1c) although the observations have been repeated with seven beam images. The magnification was $\times 800,000$, with typical exposure times of about 1 s.

In these experimental conditions, it was found that after an initial incubation period the single-crystal specimen deteriorated by the nucleation of new, small crystallites, apparently on the surfaces of the parent crystal. The deterioration time varied from one specimen to another and was typically several hours. Close examination of the specimen edge revealed that individual bright spots were moving about by a random jump process. Jumps occurred several times per second, usually by a distance equal to the atomic column separation. Not all edge spots were moving simultaneously; rather one or two would be moving back and forth along a crystal facet. Unfortunately the time both for transporting photographic

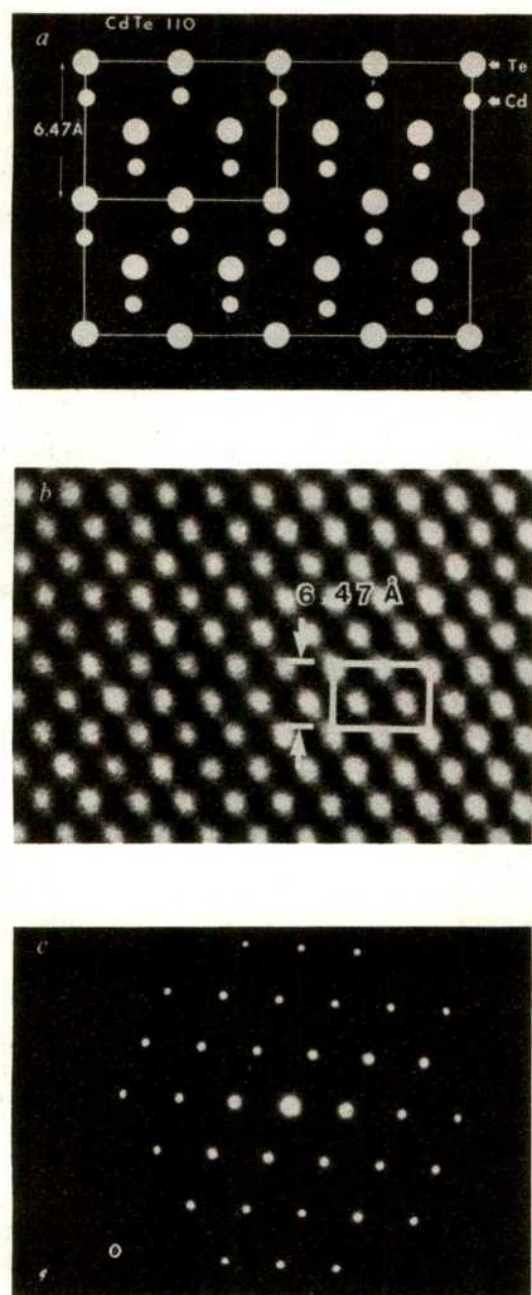


Fig. 1 *a*, Projection of the CdTe structure in the (110) orientation. *b*, Typical structure image of CdTe (atoms unresolved), with the unit cell outlined. *c*, Electron diffraction pattern showing the reflections used in image formation.

plates into position and for exposing the film was too long to allow the motion of an individual spot to be studied extensively. The jump rate could not be reduced to an acceptable level, for instance through decreasing the incident electron beam current density by defocusing the microscope condenser lens.

The motion of the surface spots is demonstrated here by the appearance of the specimen after intervals of several minutes. Thus it can be seen in Fig. 2 that the shape changes noticeably with time, although the crystal structure of the interior is always maintained. The general impression was that small facets were produced, the longer ones being parallel to {111}-type planes. The bright spots at the edge move from one lattice site to another. When they become incorporated into the crystal, they extend it and the image is indistinguishable from the perfect crystal image. These observations, therefore, seem to indicate that atomic motion is occurring along the edge of the specimen, and probably across all surfaces. Those atoms which move on

the surface normal to the electron beam will coincide exactly with the underlying atomic columns if tetrahedral coordination is to be maintained. Therefore one would not detect directly motion of these latter atoms.

To interpret the image-spot motion the relationship between the image and the CdTe crystal structure must be understood. As the image comprises bright spots, the specimen must have finite thickness. Image calculations¹ indicate that the thickness is probably in the range 40–100 Å. Furthermore, to detect a bright image spot an atomic column of at least several unit cells is necessary. Thus the spots at the specimen edge do not arise from single atoms nor from single CdTe molecules but rather from a significant column of atoms probably of dimension similar to that of the adjacent bulk crystal. The movement of spots represents coordinated motion of a whole or partial atomic column but not the jumping of single atoms. For the observation to be made, the number of atoms that move must be sufficient to produce a detectable image spot. In the present case this is thought to be a column of about five atoms (~20 Å long).

The image spots may coincide with three positions in the projected structure depending on the particular combination of the experimental variables of microscope defocus, spherical aberration coefficient and specimen thickness¹. These positions are at the open channels in the structure, at the position of one species or at the midway point between the closest Cd–Te atomic columns. (Only the first and last occur in images of Ge and Si⁶.) However, the image spot motion is independent of its exact position. Thus the jump of the Cd–Te molecule 'A' in Fig. 3 would be seen as a spot moving whether the spot coincides with the molecule, one atom, or the adjacent column. The argument also holds for other possible migration mechanisms indicated in Fig. 3. Distinction of the various mechanisms would require quantitative measurements of the intensity of particular image spots, at known focus settings, including careful image recording by a television system.

The driving force for the present behaviour is presumably reduction in total surface area. Using a simple formulation for atomic jump frequency ($\nu_0 \exp -Q/kT$), it is found either that the activation energy (Q) for the present observations is 0.8 eV if they occur at room temperature, or that the specimen temperature is ~110 °C if the activation energy is assumed to be the sublimation energy of CdTe (1.03 eV)⁷. The latter is thought to represent the present situation. Thus, the observations are only made after the specimen temperature rises sufficiently due to heating by the incident electron beam. As exposure continues, the motion becomes more rapid with more drastic specimen deterioration. Variation between specimens arises from their exact geometry and from the degree of contact between the specimens and a supporting metal grid. Complete isolation of the CdTe is found to result in extremely rapid specimen destruction. The slow response of the jump rate on reducing the beam current density indicates that the atomic motion is not assisted by direct electron–atom collisions (the electron energy is well below the threshold for Cd or Te displacement) or by ionization effects, and probably arises from a sluggish heat dissipation from the active area. A more sophisticated analysis would need to take into account the mechanism of coordinated atomic column motion.

Whether these results are confined to CdTe or represent a more general phenomenon, previously undetected, remains to be seen. Certainly the absence of a specimen contamination film in the present work is a factor contributing to the observations. CdTe has the lowest sublimation energy of the common semiconducting materials⁷. If the present ideas are correct then equivalent observations may be made on other similar solids if higher temperatures are achieved. For example, jump frequencies several times per second would be expected at 600 °C for silicon ($Q = 2.32$ eV), 340 °C for gallium arsenide ($Q = 1.63$ eV) and 260 °C for cadmium sulphide ($Q = 1.42$ eV)⁷. *In situ* specimen heating, together with lattice resolution, would be necessary to detect these phenomena which should soon be possible on commercial TEMs.

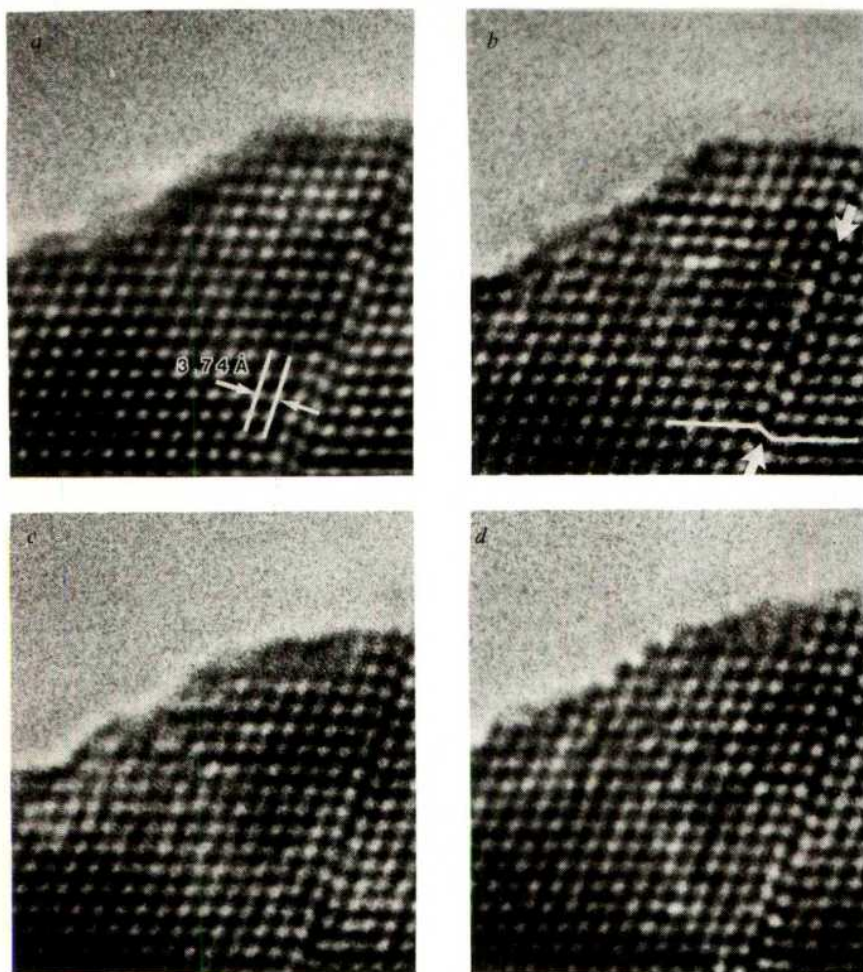


Fig. 2 Series of structure images of the same region near the specimen edge taken several minutes apart. The stacking fault (arrowed in *b*) can be used as a reference mark for each photograph. The specimen shape and its spot distribution change noticeably with time.

Direct investigations of atomic motion by electron microscopy have previously only been made without imaging the substrate. Scanning transmission microscopy results of heavy-metal atoms on amorphous carbon^{8,9} are well known, with complementary reports using fixed-beam TEM by dark field¹⁰ or bright-field imaging of atoms on graphite¹¹. Field-ion microscopy allows the imaging of both surface and substrate^{12,13} albeit of limited areas on a restricted number of materials. The present observations are thought to be the first using fixed-beam TEM, combining direct imaging of the moving species with the crystal structure of the parent material. However, unlike these other studies, they represent motion of a small group of atoms, by an undetermined mechanism, rather than of single atomic jumps.

The observations reported here have been interpreted as representing the migration of atomic columns (comprising at least five atoms and probably more) due to thermal activation induced by beam heating. A more detailed description of the

mechanism of this behaviour requires more rapid recording (for example, at television rate), extensive image calculations relating spot brightness to column thickness and a more extensive theoretical model of coordinated atomic jumps. Although this work is preliminary it has important implications. Thus with careful experimentation it should be possible to study directly self-migration on CdTe and determine the associated atomic mechanisms. If, as we suspect, the observations are not confined to CdTe, then atomic motion studies on this important class of materials are feasible. Finally, other surface-related phenomena may also be investigated, for example, the action of foreign species (such as metal atoms) and the influence of crystallography and crystal defects on the atomic behaviour. Such experiments would provide important information for understanding how the surfaces of such solids may influence their properties and their technological applications.

Financial support from the Alfred P. Sloan Foundation (R.S.), Xerox Corporation (T.Y.), Basic Energy Sciences Division of DOE (F.A.P.) and the NSF (equipment grant) is gratefully acknowledged.

Received 27 August 1980; accepted 27 January 1981.

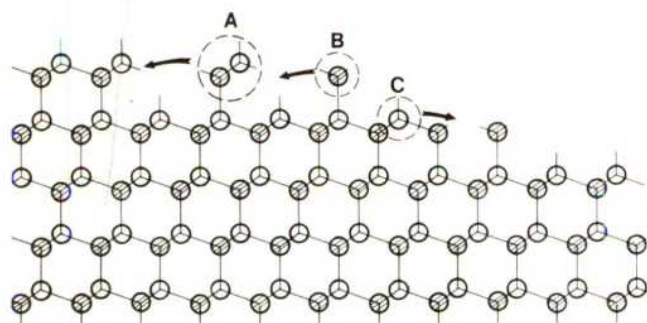


Fig. 3 Schematic model of possible mechanisms for surface atomic migration in CdTe.

1. Yamashita, T., Ponce, F. A., Pirouz, P. & Sinclair, R. (in preparation).
2. Yamashita, T. 38th *a. Proc. EMSA*, 168-169 (1980).
3. Spence, J. C. H., O'Keefe, M. A. & Kolar, H. *Optik* **49**, 307-323 (1977).
4. Rez, P. & Krivanek, O. L. *Proc. 9th Int. Congr. on Electron Microscopy* **1**, 288-289 (1978).
5. Desseaux, J., Renault, A. & Bourret, A. *Phil. Mag.* **35**, 357-372 (1977).
6. Olsen, A. & Spence, J. C. H. *Phil. Mag.* (in the press).
7. Harrison, W. A. *Electronic Structure and the Properties of Solids*, 176 (Freeman, San Francisco, 1980).
8. Isaacson, M., Ohtsuki, M. & Utlaut, M. *Introduction to Analytical Electron Microscopy* (eds Hren, J. J., Goldstein, J. I. & Joy, D. C.) 343-368 (Plenum, New York, 1979).
9. Crewe, A. V. *Chem. Scr.* **14**, 17-20 (1978-79).
10. Hashimoto, H., Yokota, Y., Takai, Y., Endoh, H. & Kumao, A. *Chem. Scr.* **14**, 125-128 (1978-79).
11. Iijima, S. *Optik* **48**, 193-214 (1977).
12. Ehrlich, G. & Hudda, F. G. *J. chem. Phys.* **44**, 1039-1049 (1966).
13. Tsong, T. T. *Chem. Scr.* **14**, 7-15 (1978-79).

High-voltage, high-resolution lattice images of dolomite

D. J. Barber

Physics Department, University of Essex, Colchester CO4 3SQ, UK

L. A. Freeman & David J. Smith

High Resolution Electron Microscope, University of Cambridge, Cambridge CB2 2QH, UK

Dolomite, $\text{CaMg}(\text{CO}_3)_2$, which belongs to the rhombohedral group of carbonates, is an important mineral commonly found in sedimentary and metamorphic environments. Ideally, it has an ordered trigonal structure ($R\bar{3}$) with planes containing atoms of only Ca, or only Mg, alternating and being interleaved with the planes containing the carbonate groups. All these planes are perpendicular to the 3-fold axis. As the microstructural deformation observed in many natural dolomites of low and medium metamorphic grade could provide an insight into the geological history, attention has recently been directed to electron microscope studies¹⁻⁴ of the mineral. Unfortunately, at a typical accelerating voltage of 100 kV, dolomite is highly susceptible to damage by the electron beam, whereas, at a voltage of 1 MV, little electron damage is noticeable¹. We report here observations of dolomite with the Cambridge University 600-kV high-resolution electron microscope^{5,6}. Although electron beam damage is still observed at the intermediate operating voltages of this microscope, it has been possible, with care, to resolve directly details of crystal lattices in two dimensions. Thus, direct confirmation has been obtained of postulated models for the common types of stacking faults observed in dolomite.

Much of the recent interest in dolomite stems from three observations:

(1) Single crystals of stoichiometric dolomite exhibit an unusual increase in mechanical strength with temperature up to $\sim 800^\circ\text{C}$ (refs 7, 2, 4), which seems to be a consequence of the detailed atomic displacements accompanying dislocation motion through the ordered structure.

(2) Attempts to synthesize dolomite in conditions similar to those in which sedimentary dolomites seem to have formed have failed^{8,9}. Moreover, although the ordered phase is more stable than disordered dolomite, the latter can exist for geologically long times. 'Geologically young' dolomites are mostly imperfectly ordered and calcium rich; embryonic analogues of the abundant, more perfect dolomites found in old environments have not been discovered.

(3) Recent work has shown that local disorder can be produced in near-perfect single crystal dolomite by slip on the (0001)^{2,4} and $\{01\bar{1}2\}$ planes⁴. X-ray structure refinements on Eugui dolomite which was thermally disordered and then quenched in a laboratory experiment indicate¹⁰, however, that other forms of disorder also occur in which some of the carbonate groups (which should ideally be oriented at 180° about the 3-fold axis with respect to the groups in adjacent carbonate planes) are misoriented by 180° . In addition, electron microscopy of non-stoichiometric sedimentary dolomites¹¹ has shown the existence of a new type of superstructure which effectively doubles the length of the a axis.

All these observations indicate the value of gaining direct information about types of disorder in dolomite. An obvious approach would involve lattice imaging techniques with a high-resolution electron microscope although, as mentioned above, dolomite is considerably electron sensitive because of local heating and ionization damage caused by the electron beam. Indeed, previous efforts at 100 kV have barely resolved the 0.54-nm (0003) lattice fringes in one-dimension only¹⁰. However, results have been improved by raising the beam

energy, which effects an appreciable decrease in the sensitivity of dolomite to beam damage, and by applying a variation of the so-called 'minimal exposure technique'¹² whereby tilting of the dolomite crystal to the required orientation is done using a focused beam of very low intensity on a small area of crystal. After focusing and correction of astigmatism, images are recorded of crystal defects and other interesting morphological features in nearby areas, previously unexposed to the intense electron beam. Local bending of the crystal lattice sometimes occurred but accurate alignment of the crystal normal and the zone axis projection was generally retained so that two-dimensional crystal structure images could be obtained.

The two-dimensional lattice image in Fig. 1 shows stacking faults produced by slip on the $\{\bar{1}012\}$ planes in a single crystal of a stoichiometric and fairly pure dolomite (Eugui) which was experimentally deformed at 20°C under a confining pressure of 0.7 GPa. The electron beam is exactly parallel to a $\langle 10\bar{1}1 \rangle$ zone axis, as indicated by the selected area diffraction pattern. Two sets of $\{\bar{1}101\}$ and one set of $\{\bar{1}012\}$ lattice planes are resolved, with spacings of 0.40 and 0.37 nm respectively.

Figure 2a shows the structure image of a thin crystal of the same Eugui dolomite which was experimentally deformed at 420°C under confining pressure in an orientation which promoted slip on the (0001) planes. In this case the electron beam is exactly parallel to a $\langle 11\bar{2}0 \rangle$ direction. $\{\bar{1}101\}$, $\{\bar{1}012\}$, $\{\bar{1}014\}$ and (0003) planes are parallel to the beam, as can be seen from Fig. 2b, which shows the atomic structure when viewed in this projection. The orientations of the carbonate groups, whose planes are strictly normal to the projection plane, are given by black triangles in this representation. A basal stacking fault, which results from slip on the (0001) $\langle 1\bar{2}10 \rangle$ macroscopic system, appears in Fig. 2a. The fault is apparently bounded by $1/3[1100]$ partial dislocations, as previously deduced from diffraction contrast experiments¹³, but it is broader in the (0001)

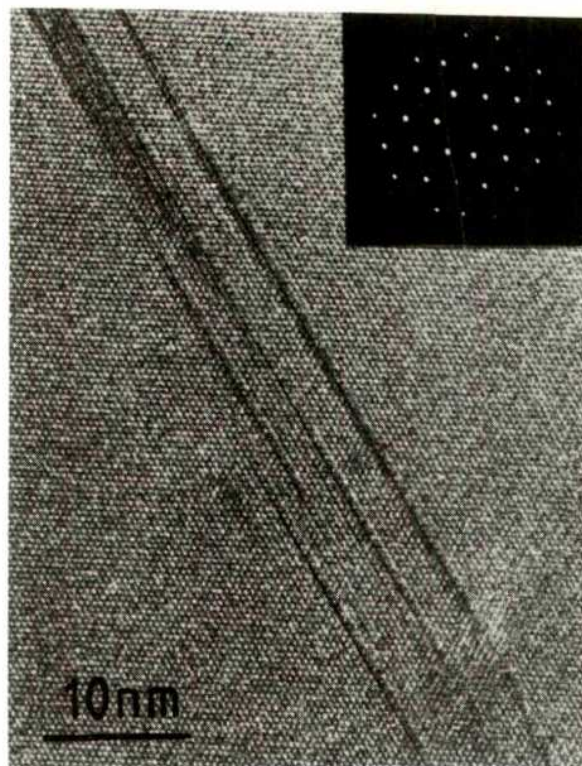


Fig. 1 Lattice image from a moderately thick crystal of dolomite in a $\langle 10\bar{1}1 \rangle$ projection showing a series of deformation-induced planar stacking faults on the $\{\bar{1}012\}$ planes, together with the corresponding electron diffraction pattern (inset). The lattice planes imaged are $\{\bar{1}011\}$, $\{\bar{1}012\}$ and $\{01\bar{1}1\}$; the faults do not produce apparent offsets of these planes.

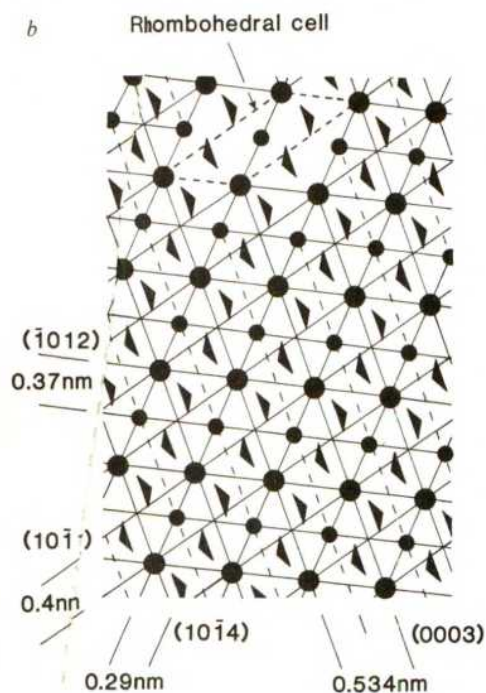
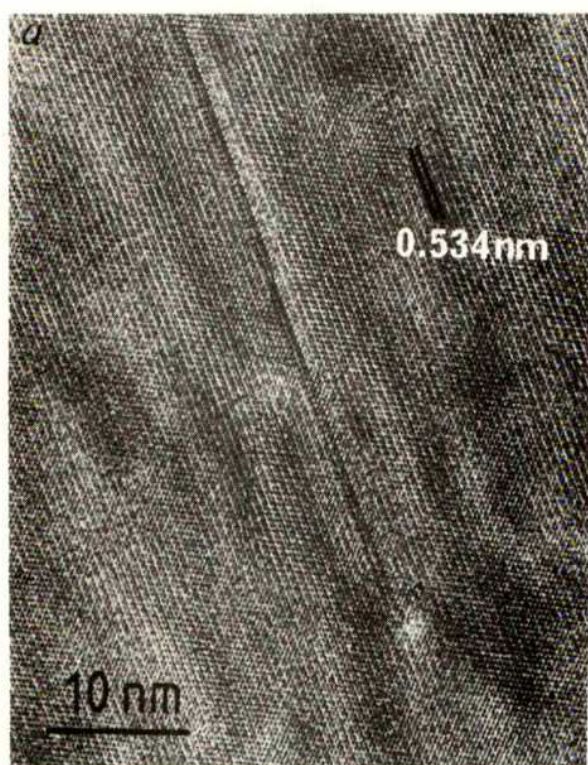


Fig. 2 *a*, Structure image of a thin crystal of dolomite taken in a $\{11\bar{2}0\}$ projection; the 0.534-nm (0003) basal plane spacing is indicated. A deformation-induced basal stacking fault is visible. Confirmation of postulated models for such stacking faults are directly obtainable. *b*, Representation of a section through a $\{11\bar{2}0\}$ plane of dolomite, showing alternating layers of Ca and Mg atoms, separated by planes of CO_3 groups parallel to the basal planes (0001). The plane of each CO_3 group is actually exactly parallel to (0001) but small (tilted) triangles are used to denote the 180° rotation between groups in adjacent planes.

direction than might be expected from elementary considerations. A more complete analysis and interpretation of (0001) and $\{10\bar{1}2\}$ stacking faults will be presented elsewhere. The disparity in cation sizes is apparent in the structure image of Fig. 2*a*, although, as might be expected, the carbonate groups are not seen in detail.

Clearly, high-voltage, high-resolution electron microscopy is a technique well suited to analysis of order/disorder problems in dolomite. Furthermore, use of this method is expected to add greatly to our knowledge of defect structures in non-stoichiometric and disordered specimens of other carbonates. Sequential images of dolomite which has undergone prolonged electron irradiation have also been obtained, from which it is believed that the nature of the electron damage process will be established. Our preliminary results suggest that the damage propagates along preferred crystallographic directions in common with earlier observations of phthalocyanines¹⁴. This damage ultimately leads to the decomposition of the carbonate to the oxides, as with calcite^{15,16}.

The Cambridge University 600-kV high-resolution electron microscope has been built as a joint project between the Cavendish Laboratory and the Department of Engineering with major financial support from the SRC. We thank the latter for continuing support, R. A. Camps for assistance with preparation of the microscope and H. C. Heard, R. J. Reeder and H. R. Wenk for provision of samples and valuable discussions.

Received 16 December 1980; accepted 18 February 1981.

1. Barber, D. J. *Tectonophysics* **39**, 193 (1977).
2. Heard, H. C., Wenk, H. R. & Barber, D. J. *EOS* **59**, 249 (1978).
3. White, J. C. & White, S. H. *Tectonophysics* **66**, 35 (1980).
4. Barber, D. J., Heard, H. C. & Wenk, H. R. *Phys. Chem. Miner.* (in the press).
5. Nixon, W. C. *et al.* in *Developments in Electron Microscopy and Analysis 1977* (ed. D. L. Misell) 13–16 (Institute of Physics, Bristol & London, 1977).
6. Cosslett, V. E. *Proc. R. Soc. A* **370**, 1 (1980).
7. Higgs, D. V. & Handin, J. W. *Bull. geol. Soc. Am.* **70**, 245 (1959).
8. Garrels, R. M., Thompson, M. E. & Siever, R. *Am. J. Sci.* **258**, 402 (1960).
9. Berner, R. A. in *Principles of Chemical Sedimentology* (McGraw-Hill, New York, 1971).
10. Reeder, R. J. thesis, Univ. California, Berkeley (1979).
11. Reeder, R. J. & Wenk, H. R. *Geophys. Res. Lett.* **6**, 77 (1979).
12. Williams, R. C. & Fisher, H. W. *J. molec. Biol.* **52**, 121 (1970).
13. Barber, D. J., Heard, H. C., Paterson, M. S. & Wenk, H. R. *Nature* **269**, 789 (1978).
14. Murata, Y., Fryer, J. R., Baird, T. & Murata, H. *Acta crystallogr.* **A33**, 198 (1977).
15. Bursage, B. J. & Pitkethly, D. R. *Phys. Status Solidi* **32**, 399 (1969).
16. Towe, K. M. *Nature* **274**, 239 (1978).

Near-surface turbulence measurements in a lake

T. M. Dillon, J. G. Richman, C. G. Hansen & M. D. Pearson

School of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

One of the most important but least understood regions of oceans and lakes is the near-surface zone. Exchanges of energy and momentum between the atmosphere and the waters below are controlled by turbulent mixing processes in the surface zone. Major advances in understanding these processes have been hampered by a lack of observations of turbulence near the surface. We have recently constructed a probe which can be used to estimate an important variable in theories of turbulence, the rate at which the kinetic energy of turbulence is dissipated, and have found a surprising though simple result: to a first approximation, scaling laws devised to describe dissipation in the atmospheric boundary layer may also be applicable in the water near the surface.

The probe is a miniaturized version of a high-frequency temperature profiler used previously in upper ocean studies¹. The probe body is a 2.5-cm diameter cylinder, 70-cm long, containing thermistor amplifiers and a pressure transducer. A toroidal buoyancy element 10 cm in diameter and 25-cm long jackets the upper probe body, and a vented drag plate 30-cm in diameter is attached to the rear. The probe is slightly buoyant in water and has a large static stability. A thermistor with a frequency response correctable to 30 Hz is mounted on the nose of the probe. Ballast is attached with a pressure-sensitive release mechanism so that after launching, the probe sinks to a pre-determined depth, releases ballast and rises to the surface at

10 cm s⁻¹. The probe is launched from a small boat, and typically rises 15–20 m away from the launch point, so that water sampled by the probe is unaffected by the boat.

Estimating the dissipation rate from high-frequency temperature measurements depends on finding the wavenumber (inverse length scale) of the smallest appreciable temperature fluctuations^{2,4}. Briefly, the method assumes a competition between the tendency of turbulent stirring to increase temperature gradients and the tendency of molecular diffusion to smooth gradients. The wavenumber at which diffusion can overcome stirring and homogenize the fluid, the Batchelor wavenumber k_B , is a function of the kinetic energy dissipation rate per unit mass, ϵ : $k_B = (\epsilon \nu^{-1} D^{-2})^{1/4}$, where ν is the kinematic viscosity and D the molecular diffusivity for temperature. Wavenumber spectra are calculated from temperature-gradient fluctuations, and by least-squares fitting to the Batchelor spectrum^{2,5,6}, k_B and hence ϵ may be found.

The validity of the method has been tested in the nepheloid layer near the ocean floor. Measurement of the stress on the ocean floor due to near-bottom currents⁷ enables the dependence of the kinetic energy dissipation rate on height above the floor to be predicted. Simultaneous measurement of the temperature microstructure was made and the Batchelor scale method was used to calculate the dissipation rate above the ocean bottom^{8,9}. Comparison of the measurement with the prediction yields excellent agreement⁴.

Measurements were made in Green Peter Reservoir, a small impoundment east of Corvallis, Oregon. The launch site had an unobstructed fetch of 3 km over water, and the depth at anchorage was 50 m. During the experiment, winds averaged 4.8 m s⁻¹. Surface waves averaged 10–20 cm in height, and occasional white-capping was observed. The temperature profile (Fig. 1) displayed an active surface mixing layer to 6 m in depth. A strong seasonal thermocline was at 7 m.

Nine casts were made through the mixing layer, and temperature gradient spectra were computed over 50-cm segments of water. From the surface to 1-m depth, the turbulence was so intense that the frequency response of the thermistor was insufficient to resolve the smallest length scales on half of the segments, and so the dissipation rate could not be reliably resolved; in these cases, we underestimate the dissipation by

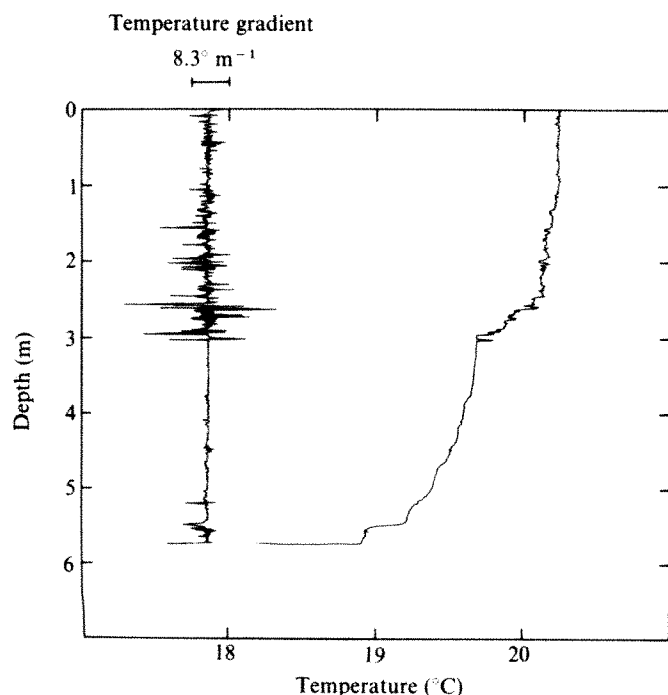


Fig. 1 Temperature (right) and temperature gradient (left) profiles for a single cast in the near-surface layer at Green Peter Reservoir, 24 July 1980. Thermocline is at 6-m depth.

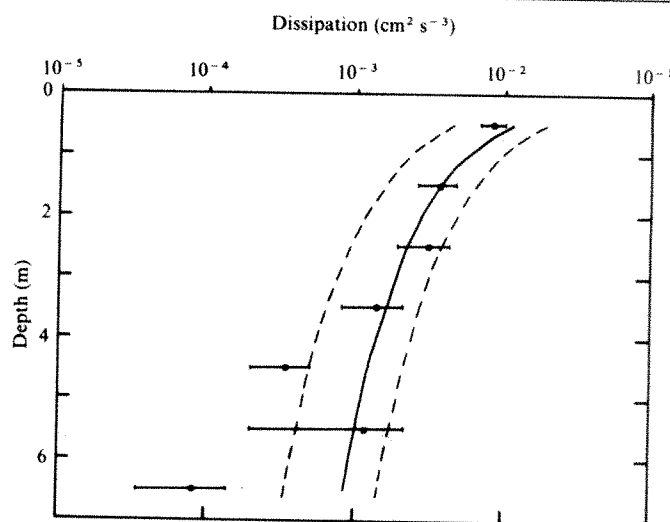


Fig. 2 Kinetic energy dissipation rate per unit mass in the near-surface layer (●). Error bars are the standard error of the mean. Solid line is $(\tau/\rho)^{3/2} \kappa^{-1} Z^{-1}$ calculated from mean winds, dotted lines are $(\tau/\rho)^{3/2} \kappa^{-1} Z^{-1}$ calculated from highest and lowest winds. Dissipation falls dramatically in the thermocline at 6 m.

setting it equal to the largest resolvable value. Below the first metre, the dissipation rate decreases, and was resolvable on 90% of the segments. At the base of the layer, the dissipation rapidly falls by an order of magnitude. Averages of the dissipation over the nine casts were calculated in 1-m intervals (Fig. 2).

No detailed theoretical framework for understanding turbulence in a wind driven surface mixed layer has yet received wide acceptance. Attempts at detailed modelling have been made¹⁰ but the complexities of possible interactions of wind, waves, breaking waves and spray, as well as the lack of near-surface observations, have kept most theories of the aquatic mixed layer at a primitive level. In contrast, the theory of the atmospheric boundary layer has been well developed¹¹. It is tempting to ignore the complex nature of the air-sea interaction and apply to the aquatic mixed layer the same framework which holds on the atmospheric side of the boundary. The atmospheric boundary layer prediction in an unstratified flow for variation of dissipation rate with distance Z from the boundary¹¹ is $\epsilon = (\tau/\rho)^{3/2} \kappa^{-1} Z^{-1}$, where κ is the von Karman constant and ρ the water density. Calculating the surface stress during our measurements by the drag law $\tau = \rho_a C_D \bar{U}^2$ (ρ_a is the air density, C_D (ref. 12) is the drag coefficient with value 1.3×10^{-3} , and $\bar{U}$ is the mean wind speed at a height of 5 m), we find remarkably close agreement with the atmospheric boundary-layer prediction. A least-squares fit to the form $\epsilon = AZ^n$ yields $n = -1.03 \pm 0.05$. Within experimental error, the dissipation is inversely proportional to the depth. Testing the magnitude of the dissipation, we find the average value $A = \epsilon Z = 0.51 \pm 0.09 \text{ cm}^3 \text{ s}^{-3}$ which is within experimental error of the independently measured value $(\tau/\rho)^{3/2} \kappa^{-1} = 0.53 \text{ cm}^3 \text{ s}^{-3}$. Exhaustive measurements of ϵ will be required to determine whether or not the atmospheric boundary layer prediction holds to more than a first approximation.

It has been suggested^{13–15} that the near-surface layer may exhibit some of the properties of a constant-stress layer, but the experimental evidence has been too fragmentary to draw firm conclusions. It is premature to suggest that the scaling will hold at all times. At large wind speeds, vigorously breaking surface waves may be an important source of turbulence, and have scaling laws different from the case of small waves and low winds.

This work was supported by the NSF under contracts OCE 7825667 and OCE 8018444, the Office of Naval Research under contract ONR 00014-79-C-0004, and the Oregon State University School of Oceanography Summer Work Program.

Received 2 September 1980; accepted 23 January 1981.

1. Caldwell, D. R., Wilcox, S. D. & Matsler, M. *Limnol. Oceanogr.* **20**, 1035-1047 (1975).
2. Dillon, T. M. & Caldwell, D. R. *J. geophys. Res.* **85**, 1910-1916 (1980).
3. Caldwell, D. R., Dillon, T. M., Brubaker, J. M., Newberger, P. A. & Paulson, C. A. *J. geophys. Res.* **85**, 1917-1924 (1980).
4. Caldwell, D. R., Chriss, T. M., Newberger, P. A. & Dillon, T. M. *J. geophys. Res.* (in the press).
5. Batchelor, G. K. *J. Fluid Mech.* **5**, 113-133 (1959).
6. Gibson, C. H. & Schwarz, W. H. *J. Fluid Mech.* **16**, 365-384 (1963).
7. Caldwell, D. R. & Chriss, T. M. *Science* **205**, 1131-1132 (1979).
8. Newberger, P. A. & Caldwell, D. R. *J. geophys. Res.* (in the press).
9. Newberger, P. A. & Caldwell, D. R. *J. geophys. Res.* (in the press).
10. Kundu, P. K. *J. phys. Oceanogr.* **10**, 220-236 (1980).
11. Lumley, J. L. & Panofsky, H. A. *The Structure of Atmospheric Turbulence* (Wiley, New York, 1964).
12. Amorcho, J. & Devries, J. J. *J. geophys. Res.* **85**, 433-442 (1980).
13. Jones, I. S. F. & Kenney, B. C. *J. geophys. Res.* **82**, 1392-1396 (1977).
14. Bye, J. A. T. *Limnol. Oceanogr.* **10**, 451-458 (1965).
15. Phillips, O. M. *The Dynamics of the Upper Ocean* 2nd edn (Cambridge University Press, 1977).

Geomagnetic induction on a transatlantic communications cable

L. V. Medford, A. Meloni,* L. J. Lanzerotti & G. P. Gregori†

Bell Laboratories, Murray Hill, New Jersey 07974, USA

The phenomenon of geomagnetically-induced currents in long conductors has been recognized since the mid-nineteenth century, when the use of the telegraph became prevalent in North America and Europe¹. Powering systems of telegraph and telephone cables are now engineered to eliminate large current surges, although the occasional significant geomagnetic disturbance, such as those produced by the February 1958², August 1972³ flare events, can still be a problem. However, induced voltages can be used, with complementary geophysical and geological data, to deduce features of the Earth's crust and upper mantle^{13,14}. Improved scientific understanding of magnetospheric phenomena means that the geomagnetic induction process in long conductors, and thus in the Earth are better understood. For example, the cable outage problem associated with the August 1972 geomagnetic storm interval was qualitatively attributed not to the 'line' currents associated with normal auroral current systems but rather to the significant enhanced currents flowing on the magnetosphere boundary at the time of an extreme compression of this boundary in greatly enhanced solar wind conditions³. We have measured the geomagnetic induction phenomenon on modern-day long cables to get better assessments of the induced currents in various types of geomagnetic variations. Our initial work was concerned with induction in the transatlantic telecommunications cable no. 6 (TAT-6) which is laid between Green Hill, Rhode Island (41.4° N, 71.7° W geographic) and St Hilaire de Riez, France (46.7° N, 1.9° W)⁴. We concentrate quantitatively here on induction by the solar-induced quiet day currents (S_q) in the ionosphere and find that calculated values agree well with observed values of S_q .

The solid line in Fig. 1 shows the approximate route of the TAT-6 cable across the North Atlantic (length 6,300 km). This cable is powered at both Green Hill and St Hilaire by constant current power supplies with opposite polarity at each end producing a current flow from west to east through the cable and repeaters^{5,6}. These repeaters are spaced approximately every 10 km along the route of the cable⁷. The cable is biased at a nominal 5,200 V positive with respect to local ground in North America and 5,200 V negative with respect to local ground at St Hilaire with a maximum capability of 7,500 V from each supply. The present measurements from the Green Hill terminal monitored the geomagnetically-induced currents in the powered

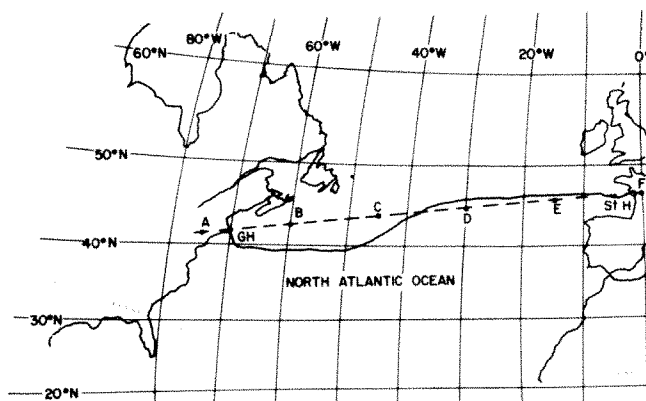


Fig. 1 Approximate location of TAT-6 cable across the North Atlantic (solid line). Model cable path (dashed line).

centre conductor by measuring changes in the conductor to local ground voltage. The voltage drop to ground at St Hilaire was not monitored in these first tests, but will be investigated in future work. To a first order, induction in the cable may be expected to produce equal and opposite voltage changes at the two ends of the system.

The constant current flow (nominal 657 mA) along the centre conductor can be pictured as coming from two sources at each cable end: one being the active power supply and the other the induced voltage produced by the rate of change with time of the geomagnetic field across the cable circuit. Thus, a measure of the time dependence of the externally-induced currents can be obtained by measuring the changes in the output voltage of the constant current power supply. The total d.c. resistance of the centre conductor and repeaters (under power) is ~ 16 k Ω . The time constant of the conductor produced by the distributed cable capacitance and the resistance (end to end) is ~ 10 s. The dashed line in Fig. 1 shows the model cable path that was taken for the modelling of the S_q field in the calculations discussed below.

In addition to monitoring the cable power supply voltage, the Earth's magnetic field in three geomagnetic components (H (north-south), D (east-west), Z (vertical)) was measured at Green Hill between 1 and 18 May 1980. The cable voltage data and the magnetometer data were individually recorded on magnetic tape using a microprocessor-based low-power digital data acquisition system with 15-bit resolution. The three components of the magnetic field and the voltage data were digitized at 2-s intervals. Timing (absolute and relative) was

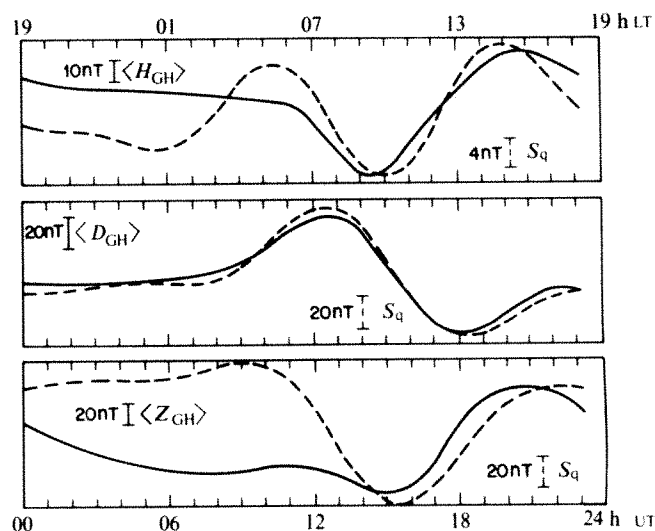


Fig. 2 Average of five quiet day magnetic field variations (solid line). Calculated magnetic field variations produced by S_q ionospheric current systems (dashed line).

Permanent addresses: * Istituto Nazionale di Geofisica, Osservatorio Geofisico Monte Porzio, 00040 Rome, Italy; † Istituto di Fisica dell' Atmosfera, CNR, 00144 Rome, Italy.

accurate in each data system to better than 1 s during the period of the measurements.

The time during which the cable measurements were made was relatively quiet geomagnetically⁸. Five of the most quiet days were selected to represent the S_q variation measured at Green Hill. These days, and their daily sum global geomagnetic activity indices ΣK_p , were: 3 May, -6; 4 May, -9; 5 May, -14; 17 May, +5; 18 May, -8. The averages of these five quiet days in the three magnetic field components measured at Green Hill are shown as solid lines in Fig. 2. These averages are obtained from the hourly mean values. The time dependences during a day show typical S_q patterns for this geomagnetic latitude, except for the local morning intervals in the H - and Z -components.

The dashed lines in Fig. 2 show the computed S_q curves expected on the basis of the spherical harmonic analysis of the potential for the daily variation by Chapman and Bartels⁹. We used the harmonic coefficients of Matsushita¹⁰ for location B on the model cable path in Fig. 1. The fit of the calculated S_q patterns using location B is quite good in the east-west component but is not good for the local morning observations in either the north-south or the vertical components. In addition, the magnitude of the S_q variation observed in the local afternoon in the north-south component is about a factor of two larger than that predicted by the model S_q calculations. These disagreements in the north-south and vertical components are

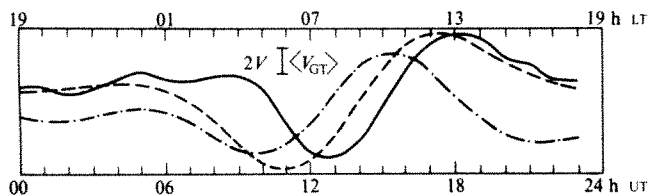


Fig. 3 Average of cable voltage variation for the five geomagnetic quiet days (solid line). Calculated induced voltage from all three magnetic components produced by S_q current system (dashed line). Calculated induced voltage from the horizontal magnetic component produced by S_q current system (dot-dashed line).

partially due to the closeness of the measurement site (~3 km) to the North Atlantic Ocean as well as to uncertainties in the spherical harmonic coefficients used in the S_q calculations. Note that data (not shown here) from our magnetometer station at Durham, New Hampshire, (44° N, 72° W geographic), located ~30–35 km from the ocean, show better agreement in the north-south component for the local morning S_q variations than do the Green Hill data.

The solid line in Fig. 3 shows the measured average variation of the cable voltage as a function of local time for the same five days as the magnetic field data of Fig. 2. This average was made using the hourly mean values and computing from them the five day average, in the same manner as was done for the S_q curves of Fig. 2. The relative expected induced voltage from induction by the S_q current system was modelled by considering the universal time (UT) change in the instantaneous S_q magnetic field pattern observed across the cable. To simplify the analysis, the S_q patterns at six locations (A–F in Fig. 1) were used for each hour of UT; the time derivative of the S_q magnetic field pattern was calculated at each of the six points. These six voltages ($V = -d\Phi(B)/dt$, where $\Phi = \mathbf{B} \cdot \mathbf{n}A$ is the magnetic flux through the surface S with area A) were averaged to represent the induced voltage across the cable at a given time. This induced voltage is of opposite polarity to the measured voltage. These calculated relative voltages (with changed signs) are superimposed on the measured voltage in Fig. 3. The dashed line is the calculated relative induction considering the total magnetic field changes from the S_q magnetic field pattern in all three components; that is, the total field contribution. The dot-dashed line is the calculated relative voltage considering only the horizontal component. The relative behaviour of this former calculated profile is a reasonably good approximation (with a 1- to 2-h

phase shift) to the observed voltage. (Induction currents from possible tidal^{2,11,13,14} and ocean current¹² effects are not considered important for this work^{2,11}.) A better agreement in terms of phase is obtained by assuming that the induction process that changes the Green Hill voltage occurs closer to the Green Hill end, rather than along the entire length of cable. Calculations of the S_q voltage pattern from points A–D (Fig. 1) show no phase shift. Physical understanding of this effect could be obtained by simultaneous measurements at both Green Hill and St Hilaire.

It is not possible to compare the magnitudes of the calculated voltage values with the observed S_q induced voltage without making some assumption as to the area of the Earth through which the induction process operates. An approximate equality between the relative calculated amplitudes and the observed induced voltage (~5 V) from the S_q pattern would imply an area of ~10¹² m² influenced by the induction.

Future work on induction by shorter-period geomagnetic variations should result in a better understanding of the geomagnetic induction phenomenon in varying geomagnetic conditions as well as at the different Earth locations over which a cable is laid.

We thank S. T. Brewer, B. H. Hamilton, and G. R. Ostby for advice and technical assistance, also G. Hailey and R. Foran for their support and help with our instrumentation at the Green Hill Terminal, and H. Wilhelmson for his careful reading of the manuscript.

Received 29 October 1980; accepted 28 January 1981.

1. Barlow, W. H. *Phil. Trans. R. Soc.* **61** (1849).
2. Axe, G. E. *Post Off. Electr. Engr. J.* **61**, 37 (1968).
3. Anderson, C. W., III, Lanzerotti, L. J. & MacLennan, C. G. *Bell Sys. Tech. J.* **53**, 1817 (1974).
4. Ehrbar, R. D., Ford, A. E. & Gerbier, G. *Bell Sys. Tech. J.* **57**, 2313 (1978).
5. Calkin, E. T., Golio, I., Schatz, W. J., Schroeder, R. E. & Shull, D. S. *Bell Sys. Tech. J.* **57**, 2497 (1978).
6. Hamilton, B. H. *IEEE Trans. Ind. Appl.* **IA-12**, 378 (1976).
7. Brewer, S. T., Easton, R. L., Soulier, H. & Taylor, S. A. *Bell Sys. Tech. J.* **57**, 2319 (1978).
8. *Solar Geophysical Data*, No. 431, Part 1 (Boulder, 1980).
9. Chapman, S. & Bartels, J. *Geomagnetism* **1**, 214 (1940).
10. Matsushita, S. in *Phys. Geomagn. Phenomena* **1**, 301 (1967).
11. Richards, M. L. *Ann. Geophys.* **33**, 177 (1977).
12. Wertheim, G. K. *Trans. Am. geophys. Union*, **35**, 872 (1954).
13. Runcorn, S. K. *Nature* **202**, 10 (1964).
14. Duffus, H. J. & Fowler, N. R. *Can. J. Earth Sci.* **11**, 873 (1974).

Palaeomagnetic determination of emplacement temperature of Vesuvius AD 79 pyroclastic deposits

Dennis V. Kent*, Dragoslav Ninkovich*,
Tullio Pescatore† & Stephen R. J. Sparks‡

*Lamont-Doherty Geological Observatory, Palisades,
New York 10964, USA

†Institute of Geology and Geophysics, University of Naples, 80138,
Naples, Italy

‡Department of Earth Sciences, University of Cambridge,
Cambridge CB2 3EW, UK

The city of Herculaneum was buried under 20 m of pyroclastic deposits during the AD 79 eruption of Vesuvius, whose crater is only 7 km to the east. These deposits have been interpreted as the deposits of mudflows¹ or hot pyroclastic flows^{2–5}. Maury's studies⁶ of incinerated wood in Herculaneum demonstrate heating to at least 400 °C. We have studied the variation of remanent magnetism with temperature for specimens taken from the deposits, including specimens of the matrix material and of embedded lithic fragments. We conclude that the temperature of the deposit at emplacement is unlikely to have been greater than 400 °C, which further supports the interpretation of the pyroclastic deposits at Herculaneum as largely ignimbrites (hot pyroclastic flow deposits).

Ten oriented drill-core samples were collected from an excavation wall in the pyroclastic flow deposits at the western end of Herculaneum. The unit sampled is typical of the deposits

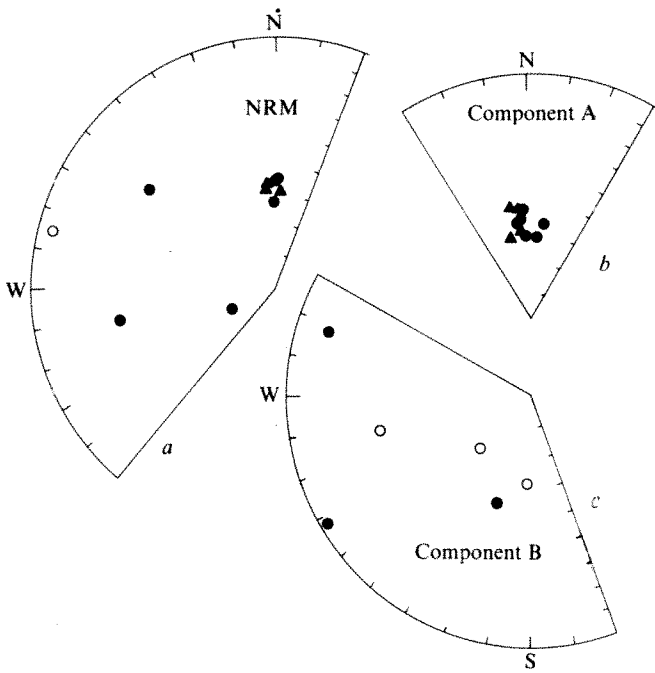


Fig. 1. Palaeomagnetic directions of pyroclastic samples from Herculaneum plotted on segments of equal-area projections. *a*, NRM directions; *b*, A components in lithic clast samples (●) and 400 °C thermally demagnetized NRM of matrix samples (▲); *c*, B components in lithic clasts (closed symbols on lower, open symbols on upper hemisphere of projection). Present geomagnetic field direction at sampling locality is dec = 359°; inc = 55°.

which bury the town and appears homogeneous and unstratified, consisting of a poorly sorted consolidated mixture of ash and pumice with local occurrences of angular lithic fragments of older volcanic rocks, limestone and brick. The volcanological stratigraphy at Herculaneum is detailed in ref. 7. The ash and pumaceous matrix is represented by samples H1 to H4; a variety of lithic fragments in addition to a large piece of light-coloured pumice, all between 10 and 20 cm in largest dimension, are represented by samples H5 to H10 (Table 1). The samples were taken over a stratigraphic interval of ~2 m, starting ~1 m above the excavation floor.

Most of the directions of the sample natural remanent magnetization (NRM) (Fig. 1*a*) group near to the present geomagnetic field for the site locality but four of the six samples of lithic fragments give divergent and scattered directions. The NRM intensities and Königsberger values (ratio of remanent magnetization to magnetization induced in present geomagnetic field) show only small variation amongst the samples of matrix but are more variable for the samples of lithic fragments (Table 1). In all samples, however, the remanent magnetizations exceed the induced components by anything between a factor of 4 and a factor of 69.

Progressive thermal demagnetization experiments show that the NRM of the the matrix samples consists of a single stable component similar in direction to the present orientation of the geomagnetic field and which persists up to demagnetization temperatures of 550 °C, when it disappears (Fig. 2*a*). In contrast, the NRM of all six lithic fragments is the result of two components of magnetization: A and B. The A component is identified in the initial stages of thermal demagnetization up to temperatures of 350–400 °C. The B component is the magnetization remaining in the final stages of thermal demagnetization, up to temperatures between 550 and 600 °C when the magnetization of the samples disappears (Fig. 2*b, c*). The directions of the A and B component (Fig. 1*b, c* and Table 1) in each lithic fragment sample are readily determined by regression analysis on the linear segments of the orthogonal vector end point demagnetization diagrams, as shown in Fig. 2. Where two specimens could be obtained for measurement from a drill-core sample of a lithic clast, both specimens provided similar magnetic characteristics.

The mean direction of the A component from the six lithic samples is dec = 359.2°, inc = 56.8°, $\alpha_{95} = 4.2^\circ$ and is statistically indistinguishable from the mean direction of the matrix samples (Table 1) or from the present geomagnetic field direction (dec = 359°, inc = 55°). Although the B components are well characterized and show good stability in individual lithic clasts, as a group they are randomly distributed at the 95% level of confidence (Table 1).

The lithic fragments evidently possessed a remanent magnetization before their incorporation in the pyroclastic deposit. This magnetization is now preserved as component B and may have originated as a thermal remanence (TRM) during initial cooling of the basalt and pumice, possibly a detrital remanence during initial deposition of the limestone, and either

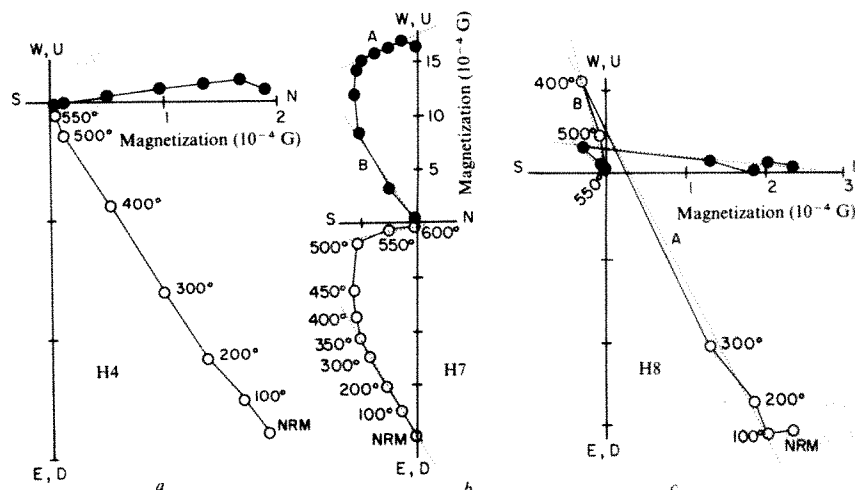
Table 1 Palaeomagnetic measurements on pyroclastic samples from Herculaneum

<i>a</i> Matrix samples		NRM				400 °C			
Sample		$J(10^{-4}\text{G})$	Q	dec	inc	dec	inc		
H1		3.4	5.8	355°	55°	351°	50°		
H2		3.2	5.9	353°	56°	349°	62°		
H3		3.1	5.2	1°	59°	355°	52°		
H4		3.4	5.6	357°	56°	355°	59°		
	Means ($n = 4$)			356.4°	56.5°	352.5°	55.8°		
				$(K = 1,020, \alpha_{95} = 2.9^\circ)$		$(K = 188, \alpha_{95} = 6.7^\circ)$			
<i>b</i> Lithic clasts		NRM				Component A		Component B	
Sample	Composition	$J(10^{-4}\text{G})$	Q	dec	inc	dec	inc	dec	inc
H5	Pumice	2.4	4.1	245°	73°	9°	56°	199°	53°
H6	Basalt	7.5	15	285°	−8	353°	55°	259°	−38°
H7	Basalt	10	14	259°	37°	356°	52°	238°	4°
H8	Limestone	3.5	69	358°	53°	6°	61°	226°	−66°
H9	Limestone	3.8	64	359°	61°	359°	61°	184°	−61°
H10	Brick(?)	15	23	309°	36°	354°	55°	286°	15°
	Means ($n = 6$)*			300.6°	49.9°	359.2°	56.8°	240.1°	−19.9°
				$(K = 4, \alpha_{95} = 37^\circ)$		$(K = 251, \alpha_{95} = 4.2^\circ)$		$(K = 2, \alpha_{95} = 59^\circ)$	
				$r = 4.80$		$r = 5.98$		$r = 3.75$	

J is intensity of magnetization; Q is Königsberger ratio; dec, inc, are declination, inclination of magnetization; means and statistics calculated using ref. 13; test for randomness from ref. 14.

*For $n = 6$, distribution is random if $r < 3.85$.

Fig. 2 Orthogonal projections of the successive end points of the magnetization vector (method from ref. 15) during progressive thermal demagnetization of representative samples from Herculaneum pyroclastic deposit. Solid symbols represent projection on the horizontal plane, open symbols on the north-south vertical plane. *a*, Matrix sample; *b*, basalt clast; *c*, limestone clast. Dashed lines in *b*, *c* are linear segments corresponding to A or B components.



a TRM or a chemical remanence when the brick was originally made. During incorporation in the pyroclastic deposit, physical rotation of the clasts resulted in the random distribution of their original (B) component magnetization.

After emplacement to their present position in the deposit, the lithic clasts became partially remagnetized, giving rise to component A which is statistically parallel in each clast. Given that the pyroclastic deposits probably experienced elevated temperatures, this magnetization is most likely a partial TRM—a supposition supported by the high Königsberger ratios. If a partial TRM mechanism is assumed for the A component, the magnetic unblocking temperatures measured in the laboratory can be used to approximate the maximum emplacement temperature from which the deposit cooled in place.

The curvature shown by some demagnetization trajectories (Fig. 2*b*) is interpreted as an overlap of blocking temperatures for magnetization components A and B. A conservative estimate for maximum unblocking temperature for component A would then correspond to the last point falling on the linear trajectory that defines this magnetization vector, which in Fig. 2*b* would be 350 °C. In other cases, (Fig. 2*c*), there is a strong inflection in the trajectories for components A and B and within the experimental uncertainties the maximum unblocking temperature for component A and the minimum unblocking temperature of the original component B coincide, at ~400 °C. The demagnetization behaviour of all lithic clast samples fell in either of these two categories and provided consistent unblocking temperature estimates for component A of between 350 and 400 °C.

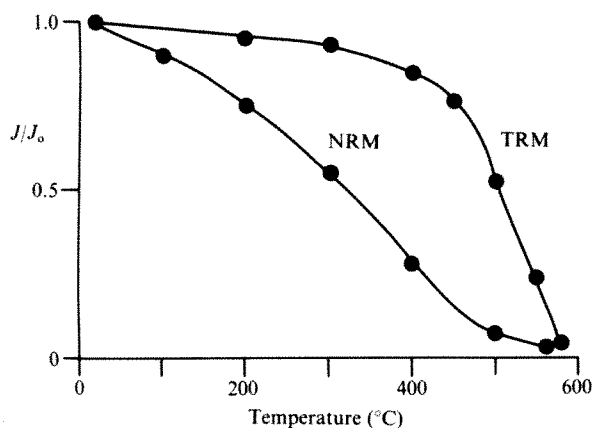


Fig. 3 Fraction of magnetization remaining with progressive thermal demagnetization for the NRM, and for a TRM produced by cooling from 600 °C in the Earth's magnetic field, of a sample of pyroclastic matrix from Herculaneum.

This consistency over the various lithologies supports acquisition of a partial TRM at a common temperature and against elevation of blocking temperatures by other processes, for example, viscous magnetization or magnetochemical alteration. On the other hand, emplacement temperatures higher than 400 °C would have more completely reset the B component magnetization.

The fact that our maximum temperature estimate coincides closely with Maury's minimum temperature estimate strongly suggests that the emplacement temperature of the AD 79 Herculaneum pyroclastic deposit sampled was indeed close to 400 °C. The deposits can therefore be considered as largely ignimbrites (hot pyroclastic flows) which in general had not attained sufficiently high temperatures to cause pervasive welding. Other evidence for a high emplacement temperature includes the occurrence of vertical lithic-rich pipes, interpreted as degassing structures, and typical fine-grained ignimbrite basal layers to individual flow units²⁻⁵.

The matrix samples give only a single magnetization component which is similar in direction to the A component of the lithic clasts although the unblocking temperatures extend to 550–660 °C, ~200 °C higher than for component A. Given the close proximity of the sampling sites for the matrix and clast samples, it is unlikely that the higher maximum unblocking temperatures observed for the matrix material reflects a correspondingly higher emplacement temperature than indicated by the clast magnetizations. The unblocking temperature spectra of the matrix samples moreover are not consistent with the spectrum of a TRM produced by cooling a matrix sample from 600 °C in the Earth's field (Fig. 3). The NRM of the matrix material is therefore more likely to contain a contribution from a magnetization besides a partial TRM from 400 °C, perhaps a detrital remanence which, from the similarity in direction, was also acquired during emplacement of the pyroclastic deposit but which is typically carried by a grain assemblage with a wide distribution of blocking temperatures. Thus while bulk samples of the matrix may provide a valid record of geomagnetic field orientation corresponding to the time the deposits are formed, they are not as such readily useful for emplacement temperature determination.

Palaeomagnetic techniques have been used to discriminate hot and cold emplacement of pyroclastic deposits, for example, in Japan⁸, Santorini⁹ and the western US¹⁰. The Herculaneum pyroclastic flow can be categorized as a type III deposit (emplaced below their maximum blocking temperature but above the ambient temperatures) of Hoblitt and Kellogg¹⁰, of which they found no example. As palaeomagnetic techniques for emplacement temperature estimation depend largely on the common occurrence of magnetizable materials in pyroclastic deposits, the techniques should be of particular value in estimating emplacement temperatures of unwelded pyroclastic flow units in which lithological evidence for heating is ambiguous, as

at Herculaneum. The size range of lithic fragments suitable for such experiments can be extended downward for example, to matrix components, by use of microanalytical techniques^{11,12} while unblocking temperature profiles through very large lithic clasts may provide additional information on the thermal history of the pyroclastic deposit. Moreover, estimates of maximum emplacement temperature by the magnetic method, which are possible up to the Curie point of the constituent magnetic mineral (for example, 585°C for magnetite) can extend the range of temperature estimates obtained from studies of organic matter, which appear to be most definitive from 100°C to ~350°C (ref. 6).

We thank Doris Lafferty for laboratory magnetic measurements. Work was supported in part by NSF grant EAR79-09985 to D.V.K. and D.N. Lamont-Doherty Geological Observatory contribution 3114.

Received 3 November 1980; accepted 3 February 1981.

1. Maiuri, A. *Scient. Am.* **198**, 68 (1958); *Herculaneum* (Rome 1977).
2. Merrill, E. T. *Am. J. Archeol.* **22**, 304 (1918); **24**, 262 (1920).
3. Sparks, R. S. J. & Walker, G. P. L. *Nature* **241**, 62 (1973).
4. Sparks, R. S. J., Self, S. & Walker, G. P. L. *Geology* **1**, 115 (1973).
5. Sparks, R. S. J. *New Scientist* **59**, 134 (1973).
6. Maury, R. C. *Bull. Centre Rech. Pan-SNPA* **10**, 289 (1976).
7. Sigurdsson, H., Cashdollar, S. & Sparks, R. S. J. *Am. J. Archeol.* (in the press).
8. Aramaki, S. & Akimoto, S. I. *Am. J. Sci.* **255**, 619 (1957).
9. Wright, J. V. *Bull. volcan.* **41**, 131 (1978).
10. Hoblitt, R. P. & Kellogg, K. S. *Bull. geol. Soc. Am.* **90**, 633 (1979).
11. Wu, Y. T., Fuller M. & Schmidt, V. A. *Earth planet. Sci. Lett.* **23**, 275 (1974).
12. Van der Voo, R. & Geissman, J. W. *EOS* **59**, 1060 (1978).
13. Fisher, R. A. *Proc. R. Soc. A* **A217**, 295 (1953).
14. Watson, G. S. *Mon. Not. R. ast. Soc., Geophys. Suppl.* **7**, 160 (1956).
15. Zijdeveld, J. D. A. in *Methods in Palaeomagnetism*, 254 (Elsevier, Amsterdam 1967).

Petrogenesis of Andean andesites from combined O-Sr isotope relationships

Russell S. Harmon

Isotope Geology Unit, Scottish Universities Research and Reactor Centre, East Kilbride G75 0QU, UK

Richard S. Thorpe & Peter W. Francis

Department of Earth Sciences, The Open University, Milton Keynes MK7 6AA, UK

The descent of the oceanic Nazca plate below the western margin of South America is responsible for the seismicity, tectonism and magmatism characteristic of Andean-type continental margins. Within this geological setting there is much debate concerning the petrogenesis of the calc-alkaline volcanic association. A thermal model for the descent of oceanic below continental lithosphere suggests that melting of subducted oceanic crust, the asthenospheric mantle wedge, and the lower crust might occur¹, and hence calc-alkaline volcanic rocks erupted at the surface in such regions might be derived from one or more of these sources. In the Andes, for example, Pichler and Zeil² have invoked partial melting of lower continental crust, while others³⁻⁵ envisage petrogenesis of andesites from sub-crustal mantle sources possibly in conjunction with crustal contamination⁶⁻⁹. We present here new O-isotope data for volcanic rocks from three contrasted volcanic settings in Ecuador, north Chile, and north-west Argentina. We then consider the significance of observed O-Sr isotope relationships, and attempt to evaluate the important factors in the petrogenesis of Andean andesitic lavas.

Active volcanism in the Andes occurs in three distinct regions: a northern zone (5°N–2°S), a central zone (16–28°S), and a southern zone (33–52°S); each having andesitic associations with distinctive chemical and isotopic compositions. The volcanic centres selected for study are in the northern and central volcanic zones. The northern zone in Ecuador and Columbia consists of composite volcanoes composed predominantly of basaltic andesite built on a 40–50-km thick sialic crust^{10,11}. Its sub-volcanic basement consists of metamorphic rocks of Cretaceous or older age¹². The central zone is characterized by composite volcanoes and large-volume ignimbrite sheets which overlie a thicker and more ancient sialic crust. Although the oldest rocks exposed adjacent to the active volcanic belt in southern Peru and northern Chile are Palaeozoic, Precambrian rocks occur to the west on the Peruvian coast¹³ and to the east below the Bolivian altiplano^{14,15} suggesting that the central volcanic belt is underlain by Precambrian rocks¹⁶. It has also been argued¹⁷ that the crust in the central Andes nearly doubled in thickness in the Cenozoic, so that the volcanoes of north Chile and south Peru rest on the central keel of the Andes where the crust is ~70 km thick. The volcanic setting in north-west Argentina is different to that to the west in southern Peru and northern Chile—here volcanic rocks occur on the 'Pampean Ranges' massif, a 'basin-and-range' province which comprises a wide variety of granitic and metamorphic rocks¹⁸ for which data indicate a late Precambrian to early Palaeozoic age¹⁹.

The Tertiary and Quaternary volcanic rocks of the Andes fall within the basaltic andesite–andesite–dacite classification of Peccerillo and Taylor²⁰. Chemical and isotopic compositions vary systematically across the Andean orogenic belt with SiO₂, K₂O and (⁸⁷Sr/⁸⁶Sr)_i increasing from west to east^{9,21-24} with calc-alkaline lavas generally enriched in K, Rb and Sr relative to P, Nb, Nd, Zr and Ti (refs 9, 25). There are, however, important inter-regional differences. The Ecuadorian lavas are olivine-pyroxene-bearing basaltic andesites with 53–61% SiO₂, 1.1–2.2% K₂O, 492–1,144 p.p.m. Sr, and 19–65 p.p.m. Rb (refs 7, 26, 27) whereas the lavas of north Chile and South Peru are predominantly pyroxene–hornblende-bearing andesites and

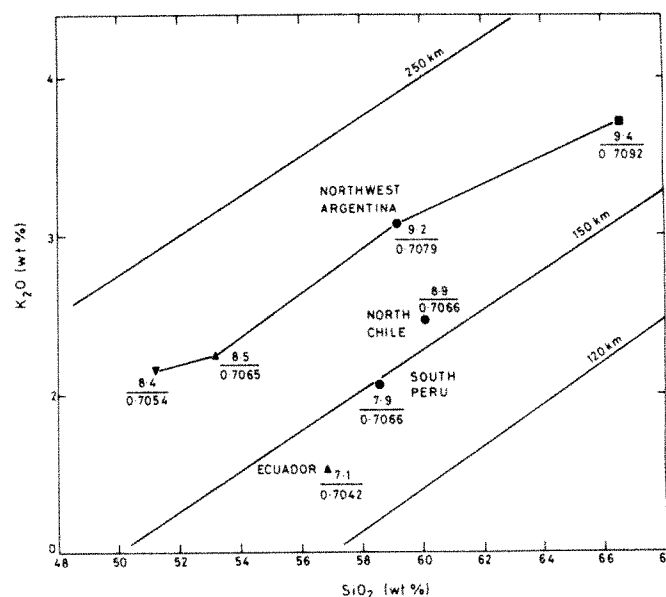


Fig. 1 Mean K₂O contents plotted against SiO₂ for volcanic centres in Ecuador, north Chile, south Peru, and north-west Argentina. The numbers shown by each data point are the average ¹⁸O/¹⁶O (upper) and ⁸⁷Sr/⁸⁶Sr (lower) ratios. The lines labelled 120, 150 and 250 km are inferred contours to the depth of the seismic zone³³. ▽, Basalt; ▲, basaltic andesite; ●, andesite; ■, dacite.

dacites with 56–66% SiO₂, 1.6–3.2% K₂O, 441–806 p.p.m. Sr, and 35–107 p.p.m. Rb (refs 17, 24). Thus, the respective mean SiO₂ and K₂O contents for the northern zone lavas are ~57 and 1.5% as compared with 60 and 2.2% for the western central zone lavas. By contrast, the Argentinean volcanic rocks cover the complete basalt–basaltic andesite–andesite–dacite spectrum. The basalt and basaltic andesites are olivine–pyroxene-bearing lavas largely from recent scoria cones and lava flows with 51–55% SiO₂, 2.0–2.7% K₂O, 477–1,048 p.p.m. Sr, and 45–75 p.p.m. Rb, whereas the andesites and dacites are pyroxene–hornblende-bearing lavas with 58–68% SiO₂, 2.9–4.2% K₂O, 236–645 p.p.m. Sr, and 52–269 p.p.m. Rb (ref. 28 and G. A. Kretzschmar, personal communication).

¹⁸O/¹⁶O ratios were determined for 33 samples from the volcanic centres of Chimborazo, Tungurahua, Sincholhua, Cotopaxi, Punalica and Calpi in Ecuador, San Pedro and San Pablo in north Chile, and Cerro Galan in north-west Argentina (Fig. 1) using the same powders for which chemical composition as well as Sr isotope and Nd isotope ratios had previously been determined. ¹⁸O/¹⁶O ratios are presented here relative to SMOW, with pertinent chemical and isotopic data: overall analytical precision is 0.1–0.2%. The combined O–Sr isotopic data are shown in Figs 1 and 2 with those of Magaritz *et al.*²⁹

All samples analysed were petrographically fresh and there is no chemical or isotopic evidence to suggest that the lavas had been subject to significant hydrothermal or deuteric alteration. Water contents are low, varying from 0.8–0.55 wt % for basalts and basaltic andesites to 0.34–1.37 wt % for andesites and dacites. Low-temperature alteration products, veining and phenocryst reaction rims are absent. However, an isotopic study of the south Peruvian andesites²⁹ showed that petrographically fresh andesitic lavas had plagioclase $\delta^{18}\text{O}$ values which were lower by an average of 0.6‰ than the corresponding whole-rock $\delta^{18}\text{O}$ values; this was attributed to a small amount of post-eruptive hydrothermal alteration. In the Ecuadorian samples plagioclase $\delta^{18}\text{O}$ values are equivalent to whole-rock $\delta^{18}\text{O}$ values, reflecting crystal–liquid equilibrium at the time of eruption.

Combined isotopic studies have proved useful in identifying the source(s) and tracing the evolution of plutonic and volcanic rocks formed at convergent plate margins. The generation of calc-alkaline magmas is a complex process which might involve partial melting of subducted oceanic crust, the overlying asthenospheric mantle wedge, and/or lower continental crust as well as subsequent fractional crystallization and contamination by upper continental crust during transit. Here the combined O–Sr isotopic data for the Andean lavas are used to comment on the petrogenesis of andesites.

The Ecuadorian lavas are the least evolved of the Andean volcanics and have the lowest ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr ratios (6.5–7.7‰; 0.7038–0.7044). The small ¹⁸O- and ⁸⁷Sr-enrichment observed in these rocks relative to typical mantle (6.0 ± 0.5‰; 0.702–0.704) could result from direct involvement of slab-released volatiles in the partial melting of overlying mantle during dehydration of the slab, by partial melting of small amounts of hydrothermally altered oceanic slab itself followed by re-equilibration in the mantle, or by partial melting of a pre-existing, alkali metasomatized mantle. Although these stable isotopic data are inadequate for a choice to be made between these alternatives, the overall chemical and isotopic characteristics of the Ecuadorian lavas suggest an alkali-enriched mantle source.

By comparison with the Ecuadorian lavas, ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr ratios are higher and more variable for the central zone volcanics in north Chile (8.0–10.8‰; 0.7059–0.7072), south Peru (7.0–8.6‰; 0.7054–0.7076), and north-west Argentina (8.3–10.7‰; 0.7054–0.7097). Because fractional crystallization cannot significantly alter the O- or Sr-isotopic composition of a magma, the observed differences between the northern and central volcanic zones must indicate a different or additional factor in the petrogenesis of the central zone lavas. If the isotopic character of the Ecuadorian lavas is representative of pristine

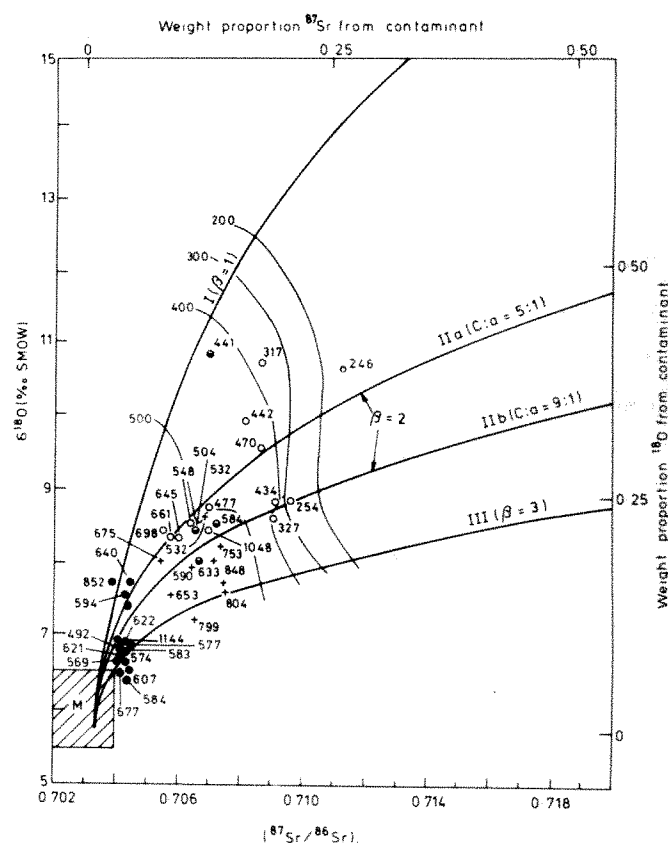


Fig. 2 Plot of $\delta^{18}\text{O}$ versus $(^{87}\text{Sr}/^{86}\text{Sr})_i$ for the Andean lavas analysed in this study. The solid curves I, IIa, IIb and III show the paths that would be followed during the process of combined assimilation-fractional crystallization for end-member compositions of hypothetical mantle ($^{18}\text{O} = 5.7\text{‰}$; $^{87}\text{Sr}/^{86}\text{Sr} = 0.703$) and an isotopically evolved crust ($^{18}\text{O} = 19\text{‰}$, $^{87}\text{Sr}/^{86}\text{Sr} = 0.735$). The curves, based on a ratio of Sr concentration in the end members of 5:1, are shown for three different (Sr cumulate/Sr melt) ratios (I, $\beta = 1$; II, $\beta = 2$ and III, $\beta = 3$) as well as two different ratios of cumulate to assimilated rocks (IIa, C:a = 9:1, IIb, C:a = 5:1) (for further discussion see ref. 34). ●, Ecuador; ●, North Chile; +, South Peru; ○, north-west Argentina. The number by each data point indicates the Sr content.

partial melts of a mantle source then those for the central zone lavas must reflect a different mantle or lower crustal source, or a contamination of magmas produced from the same mantle source.

Chemical and isotopic variations for the Andean lavas (Fig. 1) show that the K₂O–SiO₂ variations define linear trends which parallel those observed by Ninkovich and Hays³³ for island arcs and that for a given SiO₂ content there is a progressive increase in K₂O, ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr. Note that such increases occur in two directions on such a plot; enrichment of ¹⁸O and ⁸⁷Sr perpendicular to the depth contours of the seismic zone, reflecting inter-regional differences, and enrichment of ¹⁸O and ⁸⁷Sr parallel to depth contours indicating variation at a single volcano (for example, San Pedro–San Pablo, north Chile) or within a single volcanic province (for example, Cerro Galan, north-west Argentina). We interpret the former effect as indicative of source variation and the latter as contamination variation imposed on inherited source characteristics.

The most significant feature of the isotopic relationships for the Andean lavas is the positive correlation between ¹⁸O/¹⁶O and ⁸⁷Sr/⁸⁶Sr ratios (Fig. 1). Variations³⁴ of these two isotope ratios result from geologically independent processes, respectively the fractionation of stable O isotopes during the formation of low temperature sedimentary rocks at the Earth's surface and the radioactive decay of ⁸⁷Rb to ⁸⁷Sr. Thus a positive correlation

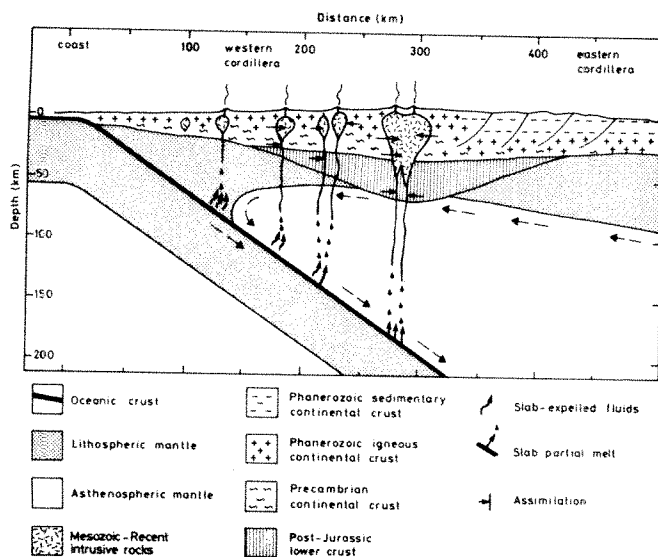


Fig. 3 The present crustal and mantle structure beneath the Andes at latitude 21°S; the known and inferred geological relationships; concepts regarding petrogenesis of Andean calc-alkaline magmas.

of ^{18}O and ^{87}Sr enrichment within a suite of genetically related igneous rocks must represent a major involvement with, or contamination by, continental crust. This correlation of O and Sr isotopic composition occurs within the Cerro Galan volcanic association of north-west Argentina, and also characterizes the overall O-Sr isotopic variation seen with the central and northern volcanic zones, although the relationship is different to those that have recently been described from two other areas of calc-alkaline magmatism, the Peninsular Range batholith of Southern California and Baja California³⁰ and the 'Newer' granites of the British Caledonides³¹. In each of these areas the positive correlations of $\delta^{18}\text{O}$ with $^{87}\text{Sr}/^{86}\text{Sr}$ extrapolate to O and Sr isotopic values representative of mantle compositions, suggesting that the calc-alkaline magmas in these three areas may have been largely derived from the mantle.

To evaluate the scale of the contamination variation, multi-component mixing models³⁴ have been used to evaluate the possible role of continental crust in the petrogenesis of the Andean lavas. Figure 2 shows the effects of a process of combined fractional crystallization-assimilation between a mantle and crystal end members with respectively low and high $^{18}\text{O}/^{16}\text{O}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios. Model lines are shown for ratios of the concentration of Sr in the cumulate material to that in the assimilated material (B) of 1 (curve I), 2 (curve II) and 3 (curve III). Curves II and IIb are respectively for bulk ratios of contaminant to assimilated material of 9:1 and 5:1. The positions of the curves in Fig. 2 depend on the exact compositions of the model end members. The low- ^{18}O , low- ^{87}Sr end member is reasonably well constrained within the limits Sr = 800–1,000 p.p.m.; $^{87}\text{Sr}/^{86}\text{Sr} = 0.702\text{--}0.704$, $^{18}\text{O}/^{16}\text{O} = 6.9 \pm 0.5$ considered characteristic of the mantle. By contrast, the composition of the high- ^{18}O , high- ^{87}Sr end member is arbitrary. The selected composition, Sr = 50–200 p.p.m.; $^{87}\text{Sr}/^{86}\text{Sr} = -0.735$, and $^{18}\text{O}/^{16}\text{O} = -19\%$ may be representative of the metamorphic and igneous rocks believed to be an important constituent of the central Andean Precambrian crust. However, the choice of a slightly different end-member composition would not change the shape of the curves, only the scale of the axes, and so require a larger or smaller proportion of crustal component to explain a given $^{18}\text{O}/^{16}\text{O}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ value.

Figure 2 shows that the overall O-Sr isotopic trend for the Andean volcanic arc may be explained in terms of a process of combined assimilation and fractional crystallization between a

primitive, mantle-derived magma and isotopically evolved continental crust. This is best illustrated for the Cerro Galan volcanic suite of north-west Argentina. Francis *et al.*²⁸ considered two models for the Sr isotopic relationships (closed system bulk contamination and open system fractional crystallization-selective contamination) and concluded that the Sr-isotope variations were best explained by a small degree of selective Sr contamination of a magma ranging from basic to acidic in composition as the result of fractional crystallization. Figure 2 shows that the variation in $^{18}\text{O}/^{16}\text{O}$ ratios, 8.3–10.7%, together with an increase in $^{87}\text{Sr}/^{86}\text{Sr}$ ratio, 0.7054–0.7097, is consistent with a model involving fractional crystallization of a magma with the O- and Sr-isotopic characteristics of the least evolved sample (G49: $^{18}\text{O}/^{16}\text{O} = 8.4\%$; $^{87}\text{Sr}/^{86}\text{Sr} = 0.7054$) allied with up to 20% bulk assimilation. Although a well-defined O-Sr isotopic correlation is not present within the north Chilean and south Peruvian volcanics when considered individually, we feel this is likely to be due to the restricted compositional range of these lavas. However, it is possible that these variations might reflect primary characteristics of the source³², inherited characteristics of a source modified by crustal interaction³⁴ or the mixing of mantle and crustal magmas³⁵. However, the concomitant enrichment of ^{18}O and ^{87}Sr from the northern to central volcanic zones strongly indicates a crustal component to the central zone lavas.

As predicted by the assimilation fractional crystallization model, there is a general change in the major element chemistry coincident with the O-Sr isotopic variations (Fig. 1) as well as a general decrease in Sr content with increasing ^{18}O and ^{87}Sr content (Fig. 2). Figure 2 shows that most analyses cluster about the curves corresponding to a $\text{Sr}_{\text{cumulate}}/\text{Sr}_{\text{melt}}$ ratio of 2:1 and a cumulate to assimilate ratio of between 5:1 and 9:1. Therefore, if the assimilation-fractional crystallization process suggested by the O-Sr isotopic variations is valid, then the modelling implies that a large amount of cumulate must have formed during the evolution of the andesites and dacites from their basic precursors. Such cumulates might form parts of the batholithic intrusions inferred to be present below the Andes or might be present at depth within the lower crust. This suggestion is consistent with the recent estimate of the relative volumes of intrusive to extrusive rocks in the central Andes of Thorpe *et al.*³⁶, based on volumetric data.

Magaritz *et al.*²⁹ have recognized the necessity for an ^{18}O -enriched crustal component in the generation of andesitic lavas as a result of their O-Sr isotopic study of the Arequipa and Barroso volcanics of south Peru. After considering possible sources and by analogy with the Banda Arc, they concluded that the only reasonable source for the high ^{18}O component was continentally derived sediment which had been subducted and become involved in partial melting at depth. This argument was largely based on the absence of the $^{18}\text{O}/^{16}\text{O}$ – $^{87}\text{Sr}/^{86}\text{Sr}$ correlation which we have shown here to exist overall across the Andes. Even disregarding physical²⁸ and chemical²⁵ arguments against the involvement of subducted sediment in Andean calc-alkaline magma genesis, the isotopic data do not support such an idea. Certainly on mass balance considerations it is not feasible, given any reasonable proportions of components, to melt partially subducted-geosynclinal sediments ($\delta^{18}\text{O} \sim 13\%$; $^{87}\text{Sr}/^{86}\text{Sr} = 0.709$) together with altered oceanic crust ($\delta^{18}\text{O} \sim 7.9\%$; $^{87}\text{Sr}/^{86}\text{Sr} = 0.709$) and mantle to produce the ^{18}O - and ^{87}Sr -enriched lavas that occur in the central volcanic zone ($\delta^{18}\text{O} \sim 8\text{--}9\%$; $^{87}\text{Sr}/^{86}\text{Sr} \leq 0.711$). Likewise, the recent Pb-isotope data of Barreiro and Tilton³⁷ on the Arequipa and Barroso volcanics which are characterized by low $^{206}\text{Pb}/^{204}\text{Pb}$ ratios argue against contamination by subducted sediments—a crustal source is favoured.

The Andes are a classic setting in which to study the petrogenesis of calc-alkaline magmatism at destructive plate margins. Here the importance of subduction zone, mantle and crustal processes can be evaluated by chemical and isotopic techniques. The present data suggest an origin for andesitic magmas, illustrated in Fig. 3, which involves dehydration of the subducted

oceanic crust at depths of 60–100 km and alkali enrichment of the overlying asthenospheric mantle by LIL-enriched volatiles. The partial melting of this mantle then produces a basaltic melt slightly enriched in K, Rb and Sr relative to Zr and Ti and the REE as well as enriched in $^{87}\text{Sr}/^{86}\text{Sr}$ relative to $^{143}\text{Nd}/^{144}\text{Nd}$ and enriched in ^{18}O relative to normal mantle. This may be viewed as either a single-stage process in which alkali element enrichment occurs coincident with partial melting of primitive, homogeneous mantle, or as a two-stage process where a previously metasomatized mantle is partially melted to produce a calc-alkaline magma. Subsequent segregation of and diapiric uprise into the crust are followed by fractional crystallization and assimilation within the crust before eruption at the Earth's surface. In areas of relatively thin crust, such as in Ecuador, magmas rise to the surface as basaltic andesites largely unmodified by crustal contamination. However, in areas of thicker crust, such as in the central volcanic zone, the rising magmas are modified by a process of fractional crystallization allied with assimilation of isotopically evolved continental crust to produce andesites and dacites enriched in both ^{18}O and ^{87}Sr . Away from the subduction zone where partial melting occurs at depth in excess of 150 km, such as in north-west Argentina, magmas are probably derived from partial melting of the transformed oceanic slab with little or no volatile component so that the initial melts are enriched in both ^{18}O and ^{87}Sr relative to those magmas produced nearer the subduction zone from partial melting of the mantle. Here again slow rise of the basic melt through the thick crustal keel of the Central Andes provides an extended opportunity for fractional crystallization and crustal assimilation.

Isotopic work at the Scottish Universities Research and Reactor Centre and field work in the Andes by the Open University is supported by grants from the NERC. We thank G. A. Kretzschmar for unpublished K_2O data for Cerro Galan, J. Borthwick and J. Jocelyn for technical assistance, and C. J. Hawkesworth, H. P. Schwarcz and A. N. Halliday for helpful discussions. Miss J. McLean, Mrs E. Yuill, Mrs J. Bennett and D. MacLean helped to prepare the manuscript.

Received 26 August 1980; accepted 3 February 1981.

1. Anderson, R. N., Delong, S. E. & Schwarz, W. M. *J. Geol.* **88**, 445–451 (1980).
2. Pichler, W. & Zeil, W. *Bull. volcan.* **35**, 424–452 (1972).
3. Noble, D. C., et al. *Geology* **3**, 501–504 (1975).
4. James, D. E., Books, C. & Cuyubamba, A. *Bull. geol. Soc. Am.* **87**, 592–600 (1976).
5. Brooks, C., James, D. E. & Hart, S. R. *Science* **193**, 1086–1089 (1976).
6. Thorpe, R. S., Potts, P. J. & Francis, P. W. *Contr. Miner. Petrol.* **54**, 65–78 (1976).
7. Francis, P. W., Moorbath, S. & Thorpe, R. S. *Earth planet. Sci. Lett.* **37**, 197–202 (1977).
8. Dostal, J., Zentilli, M., Caelles, J. C. & Clark, A. H. *Contr. Miner. Petrol.* **63**, 113–128 (1977).
9. Hawkesworth, C. J. et al. *Earth planet. Sci. Lett.* **42**, 45–57 (1979).
10. Case, J. E., Barnes, J., Gabriel Paris, Q., Gonzalez, I. & Vina, A. *Bull. geol. Soc. Am.* **84**, 2895–2904 (1973).
11. Meissner, R. O., Fleuh, E. R., Stibane, F. & Berg, E. *Tectonophysics* **35**, 115–136 (1976).
12. Henderson, W. G. *J. geol. Soc. Lond.* **136**, 367–378 (1979).
13. Shackleton, R. M., Ries, A. C., Coward, M. P. & Cobbold, P. R. *J. geol. Soc. Lond.* **136**, 195–214 (1979).
14. Lehmann, A. *Geol. Rdsch* **67**, 270–285 (1978).
15. Cobbing, E. J., Ozard, J. M. & Snelling, N. J. *Bull. geol. Soc. Am.* **88**, 241–246 (1977).
16. Cobbing, E. J. & Pitcher, W. S. *Nature* **246**, 51–53 (1972).
17. James, D. E. *Bull. geol. Soc. Am.* **82**, 3325–3346 (1971).
18. Rapela, G. W. & Shaw, D. M. *Geochim. cosmochim. Acta* **43**, 1117–1129 (1979).
19. Halpern, M. & Lattore, C. O. *Rev. Ass. Geol. Argentina* **28**, 195–205 (1973).
20. Peccerillo, A. & Taylor, S. R. *Contr. Miner. Petrol.* **58**, 63–81 (1976).
21. McNutt, R. H. et al. *Earth planet. Sci. Lett.* **27**, 305–323 (1975).
22. Lefevre, C. *Contr. Miner. Petrol.* **41**, 259–272 (1973).
23. Dupuy, C. & Lefevre, C. *Contr. Miner. Petrol.* **46**, 147–157 (1974).
24. Francis, P. W., Robbol, M. J., Walker, G. P. L., Cobbold, P. R. & Coward, M. *Geol. Rdsch* **63**, 357–388 (1974).
25. Klerkx, J., Deutsch, S., Pichler, W. & Zeil, W. *J. Volcan. Geotherm. Res.* **2**, 49–71 (1977).
26. Pichler, W., Stibane, F. R. & Weyl, R. *Neues Jb. Geol. Paläontol.* **2**, 102–126 (1977).
27. Thorpe, R. S. & Francis, P. W. *Tectonophysics* **57**, 53–70 (1979).
28. Francis, P. W., Thorpe, R. S., Moorbath, S., Kretzschmar, G. A. & Hammill, M. *Earth planet. Sci. Lett.* **48**, 257–267 (1980).
29. Magaritz, M., Whitford, D. J. & James, D. E. *Earth planet. Sci. Lett.* **40**, 220–230 (1978).
30. Taylor, H. P. & Silver, L. T. *U.S. geol. Surv. Open-File Rep.* **78**–701, 423–426 (1978).
31. Harmon, R. S. & Halliday, A. N. *Nature* **283**, 21–25 (1980).
32. Hamilton, P. J., O'Nions, R. K. & Pankhurst, R. J. *Nature* **287**, 279–284 (1980).
33. Ninkovich, D. & Hays, J. D. *Earth planet. Sci. Lett.* **16**, 331–345 (1975).
34. Taylor, H. P. *Earth planet. Sci. Lett.* **47**, 243–254 (1980).
35. Halliday, A. N., Stephens, W. E. & Harmon, R. S. *J. geol. Soc. Lond.* **137**, 329–348 (1980).
36. Thorpe, R. S., Francis, P. W. & Harmon, R. S. *Proc. R. Soc. Meet. on Origin and Evolution of the Earth's Crust* (Royal Society, London, in the press).
37. Barreiro, B. & Tilton, G. R. *Proc. 26th int. geol. Congr.* **1**, 14 (1980).

Thin-skinned tectonics in the northern Rhenish Massif, Germany

R. Meissner & H. Bartelsen

Institut für Geophysik, Universität Kiel, 23 Kiel, FRG

H. Murawski

Geol.-Paläontol. Institut, Universität Frankfurt, 6 Frankfurt/Main, FRG

A network of profiles in the Hercynian mountain system (Geotransverse Rhenoharzynikum) is needed to determine details of the geological structure and find the reasons for tectonic movements. Outstanding features within the mountain system such as the large fault zone along the southern boundary of the Hunsrück¹ and the left flank of the Rhinegraben have already been investigated by reflection seismology^{2,3}. We describe here a new reflection–refraction experiment carried out in 1978 in the most northern part of the Rhenish Massif, across the Stavelot–Venn Massif near Aachen (Fig. 1). While layers in the lower crust dip to the NNW, a strong upper reflector at 3–4 km depth shows a slight dip to the SSE and indicates a prominent fault zone near Aachen, known in Belgium as the Faille du Midi. Here, a comparable and apparently identical reflector was found some 100 and 175 km to the west of our profile. We interpreted this reflector as a prominent and well lubricated thrust fault along which a huge horizontal nappe displacement took place during the last stages of the Variscan orogeny.

A special arrangement of shot points and geophone patterns was used to investigate the complicated tectonics in the northern part of the Rhenish Massif. Each of the 15 shots for the 31-km long line was recorded by three synchronized 48-channel recording units, each set out at 7.5-km length spreads, leaving two gaps of ~4 km each which were filled by individual reflection–refraction stations (see Fig. 2). Explosive charges of 50–400 kg were fired at points along the line and at their extension to the SSE, providing a total length of 37 km in the subsurface with a six-fold coverage for the main part of the profile. This simultaneous recording of such a long geophone spread provided us with huge move-out times, that is with the possibility of determining the velocity–depth structure with a very high accuracy from the reflection move-out. It was also a powerful method of determining the (predominantly horizontal) velocity from the evaluation of the first (refracted) arrivals which appeared as 12 overlapping refraction branches for the sixfold reflection coverage, hence providing sufficient redundancy for calculating near-surface velocities and corrections. (Details of this method and the interpretation work will be given elsewhere.) Figure 3 shows a stacked section of the upper 3 s two-way travel time involving the upper reflector. Strongly varying velocities within the steeply dipping surface layers are responsible for some remaining apparent undulations of the reflector. To determine its nature the following conclusions can be made.

The strong, high-frequency reflector indicates a considerable jump in impedance and hence in seismic velocity. The velocity cannot jump to a thick layer with a higher velocity because the first arrivals do not show any considerable velocity increase up to distances of 45 km, nor can the velocity jump to a thick layer with a lower velocity because no break in the first arrivals is observed. Hence, the strong reflectivity must result from a very thin layer of the order of a fraction of the dominant wavelength. Its thickness, estimated to be between 40 and 100 m, agrees both with the continuity of the first arrivals and with the single-wavelet, good quality of the reflector. A thin layer of higher velocity seems improbable because at least some high-frequency

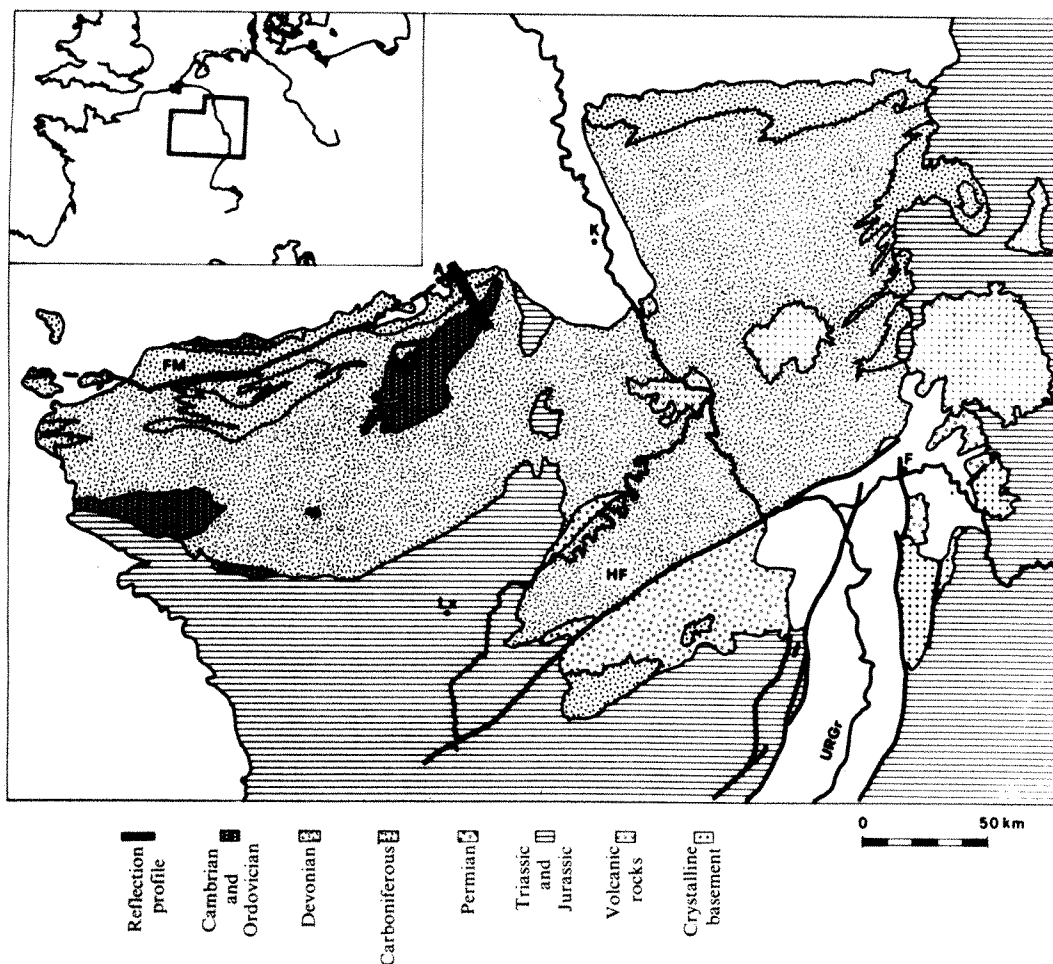


Fig. 1 Geological map of the Rhenish Massif and surrounding areas with location of profile. A, Aachen; F, Frankfurt; K, Köln; Lx, Luxembourg; FM, Faille du Midi; URGr, Upper Rhine graben; HF, Hunsrück fault.

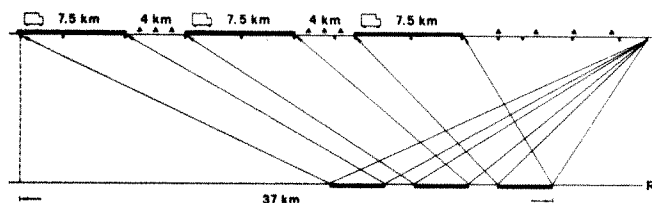


Fig. 2 Profile field set-up. ●, Shotpoints; ▲, additional wide-angle stations. Hatching on the schematic reflector R shows the 'subsurface coverage' for the first shotpoint, corresponding to the geophone spread of three recording units. Signals from all shotpoints were recorded by the geophone spreads resulting in a sixfold coverage of reflectors.

events with higher velocities would be expected in this case. The only consistent solution, therefore, is provided by a thin layer with lower velocities. Such a layer, together with the shape and the direction of the reflector which points directly towards a prominent surface fault system in the Hohes Venn area, strongly suggests that the reflecting interface resembles a thrust fault with

indications of large nappe displacement to the north-west. Figure 4 shows a geological cross-section of the upper part of the area under investigation, containing the reflector as its lower boundary.

The interpretation of the reflector as a prominent thrust fault is supported by petrological evidence, seismological and geodetic observations and by other seismic reflection results. Thrust faults, for example, generally exhibit high pore pressure, grain diminution and generally, a hydrolytic weakening⁴. Although these effects lead to a decrease in velocities^{5,6}, a high pore pressure which considerably reduces the effective pressure might have the strongest effect on the velocity decrease within a thin zone of deformation⁷. Hence it seems important that the fault zone is still seismically active in the correct overthrust direction as seen from a fault plane solution of a quake in the thrust fault area⁸. Also, an increased vertical uplift has been observed in the Hohes Venn area⁹. Today's stress pattern¹⁰ seems to be similar to that during late Variscan time, hence probably reactivating the old fault zones. Note that reflection experiments of the Famenne area, ~100 km WSW of the

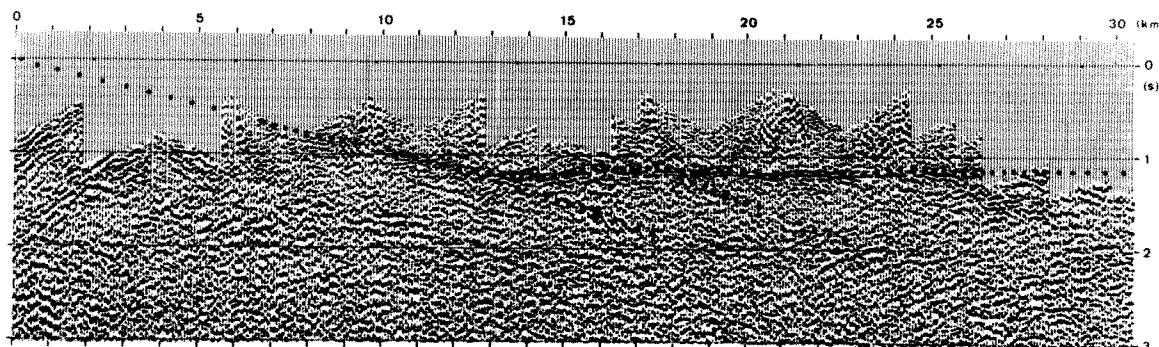


Fig. 3 Stacked section of the upper 3 s with reflector at 1.2 s two-way travel time.

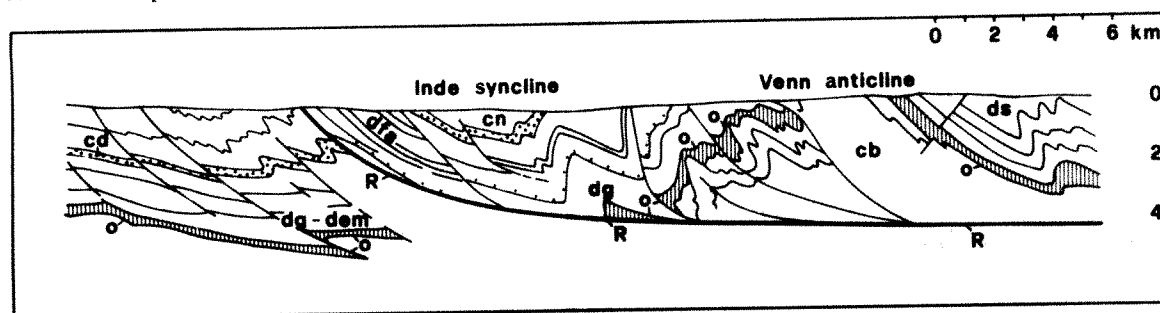


Fig. 4 Geological cross-section along the profile. R, Seismic reflector; cb, Cambrian; o, Ordovician; dg, Gedinnian; ds, Siegenian; dem, Emsian; dfa, Famennian. Carboniferous: cd, Dinantian; cn, Namurian.

present profile, have also revealed a prominent reflector at roughly the same depth as our presumed thrust fault and also dipping gently to the SSE. A comparable reflector has also been found some 175 km to the west¹² and even farther. Figure 5 shows the approximate depth of this reflector in the area of investigation. The identified reflector, interpreted as a thrust fault, covers a large area and may even extend laterally to the east and west and to greater depths to the south where no seismic investigation has been carried out.

Recent investigations in the southern Appalachians by the COCORP group¹³ and in the northern Appalachians have shown similar pronounced, shallow and slightly inclined reflectors. In the southern part they are interpreted as thrust faults on which the crystalline basement of the Appalachians has been thrust at least 260 km towards the North American continent¹³⁻¹⁵.

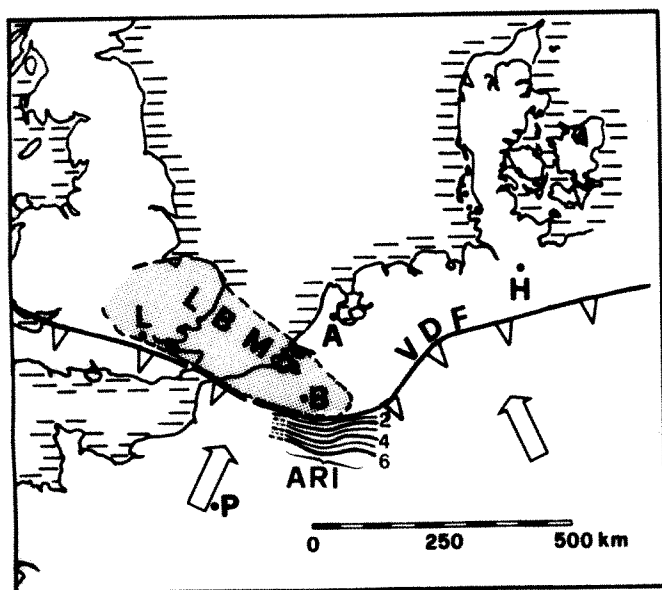


Fig. 5 Area of reflection investigation (ARI) within the Variscan deformation front (VDF). ARI shows isolines of depth of reflecting interface (thrust fault) in km. LBM, London Brabant massif; H, Hamburg; L, London; A, Amsterdam; B, Bruxelles; P, Paris.

We do not know whether the thin-skinned nappe in our test area in the Variscides was transported over a comparably large distance, but this possibility must be considered. The lateral dimensions of the nappe of at least 200 km—a lower limit, only marked by the available seismic reflection observations—indicates that large and regional compressive forces were acting during the last phases of the Variscan orogeny. We also do not know whether a pattern of approaching and colliding island arcs and microcontinents, as apparently involved in the creation of the southern Appalachians¹⁵, was also responsible for the formation of the complex Variscan tectonics. Certainly, here the deformation front also moved from SW towards the NW—towards a large continental area.

Extensive nappe displacements are well known from virtually all orogenic belts from Caledonian to recent compressional regimes, for example, Alpine-Himalayan; but in most cases the depth, the dimension and the structure of them remain unknown. For parts of the Appalachians^{14,15}, and now also for parts of the northwestern Variscides the character and the depths of the thrust planes have been revealed and their nature analysed from reflection seismology. The formation of thin-skinned nappes riding over a rather undeformed subsurface along plane, well lubricated thrust faults may indicate a final stage of compressional tectonics.

We thank P. Ziegler, for helpful discussions and our colleagues of the Geological Institute of the RTHW Aachen for the geological profile.

Received 14 October 1980; accepted 3 February 1981.

1. Ziegler, P. A. *Geologie Mijnb.* **57**, 589–626 (1978).
2. Ahorner, L. & Murawski H. *Z. ges. Geol.* **126**, 63–82 (1975).
3. Meissner, R., Bartelsen, H. & Murawski, H. *Tectonophysics* **64**, 59–84 (1980).
4. Etheridge, M. A. & Wilkie, J. C. *Tectonophysics* **58**, 159–178 (1979).
5. Kern, H. *Tectonophysics* **44**, 185–203 (1978).
6. Meissner, R. & Fakimi, M. *Geophys. J. R. astr. Soc.* **49**, 133–143 (1977).
7. Christensen, N. J. *J. Geophys. Res.* **84**, 6849–6858 (1979).
8. Ahorner, L. *Protokoll 3. DFG Kolloq.*, Neustadt 40–42 (1978).
9. Teichmüller M. & Teichmüller, R. *Fortschr. Geol. Rheinld Westf.* **27**, 323–355 (1979).
10. Illies, J. H. & Greiner, G. *Tectonophysics* **52**, 349–359 (1979).
11. Bless, M. J. M., Bouckaert, J. & Paproth, E. *Meded. Rijks geol. Dienst* **32-1**, 3–13 (1980).
12. Clément, J. *Ann. Soc. géol. Nord* **83**, 237–241 (1963).
13. Brewer, J. A. & Oliver, J. E. *Rev. Earth planet. Sci.* **8**, 205–230 (1980).
14. Cook, F. A. *et al. Geology* **7**, 563–567 (1979).
15. Cook, F. A., Brown, L. D. & Oliver, J. E. *Scient. Am.* **243**, 124–138 (1980).

Group selection is implicated in the evolution of female-biased sex ratios

Robert K. Colwell

Department of Zoology, University of California, Berkeley, California 94720, USA

Female-biased sex ratios are common among arthropods that mate in small groups with episodic dispersal¹⁻³. When breeding groups are founded by more than one female, individual (within-group) selection favours an unbiased (1:1) sex ratio, regardless of group size and degree of inbreeding (or sibmating), and regardless of the intensity of local competition among siblings for mates. On the other hand, female-biased sex ratios are favoured by group selection in randomly structured populations. Existing treatments of the evolution of female-biased sex ratios^{1,2,4-6} correctly predict the result of evolution, but do not make clear that it is a balance between opposing forces at two levels of selection³. I now show this to be so.

Fisher⁷ pointed out that in a random-mating population with a biased sex ratio, an individual of the rarer sex leaves more progeny, on average, than an individual of the commoner sex, so that frequency-dependent selection favours equal investment in sons and daughters. Hamilton¹ modelled the evolution of

Table 1 Model for the evolution of female-biased sex ratios as a balance between the opposing forces of individual and group selection

For all breeding groups, let

- n = number of inseminated females founding each group
 a, b = proportion of sons among the progeny of each A-type (a) or B-type (b) female
 r = number of progeny produced by each female, regardless of type
 $\bar{p}$ = proportion of founding females that are A-type in the global population
 p' = proportion of grandmaternal gene copies that are due to A-type founding females in the global population

I Individual selection (within groups)

In a group founded by n inseminated females, of which $x = p_x n$ are A-type (each producing ar sons) and $(1 - p_x)n$ are B-type (each producing br sons), the numbers of sons and daughters among the group's progeny (F_1) are

$$s_x = n[p_x a + (1 - p_x)b] \quad \text{and} \quad d_x = n[p_x(1 - a) + (1 - p_x)(1 - b)] \quad (1)$$

Now the F_1 mate at random within the group, with mating probabilities

$$\begin{aligned} A\delta \times A\varphi: M_{AA} &= p_x^2 \left(\frac{a}{s_x} \right) \left(\frac{1 - a}{d_x} \right) r^2 n^2 \\ A\delta \times B\varphi: M_{AB} &= p_x(1 - p_x) \left(\frac{a}{s_x} \right) \left(\frac{1 - b}{d_x} \right) r^2 n^2 \\ B\delta \times A\varphi: M_{BA} &= (1 - p_x)p_x \left(\frac{b}{s_x} \right) \left(\frac{1 - a}{d_x} \right) r^2 n^2 \\ B\delta \times B\varphi: M_{BB} &= (1 - p_x)^2 \left(\frac{b}{s_x} \right) \left(\frac{1 - b}{d_x} \right) r^2 n^2 \end{aligned} \quad (2)$$

$$M_{AA} + M_{AB} + M_{BA} + M_{BB} = 1$$

where $A\delta$ means a son of an A-type founder, $B\varphi$ a daughter of a B-type founder, and so on. The proportion of grandmaternal gene copies among the F_2 that are due to A-type founders is

$$p'_x = M_{AA} + \frac{1}{2}(M_{AB} + M_{BA}) = \frac{p_x^2 a(1 - a) + \frac{1}{2}p_x(1 - p_x)[a(1 - b) + b(1 - a)]}{[p_x a + (1 - p_x)b][p_x(1 - a) + (1 - p_x)(1 - b)]} \quad (3)$$

and the total number of F_2 gene copies due to A-type founders is

$$p'_x V_x = p'_x r d_x \quad (4)$$

assuming diploidy.

For a particular breeding group, let

- x = number of founding females that are A-type (the rest, $[n - x]$, are B-type)
 $p_x = x/n$ = proportion founding females that are A-type
 p'_x = proportion of grandmaternal gene copies that are due to A-type founding females
 s_x, d_x = total number of sons (s_x) or of daughters (d_x) produced by all n founding females
 $V_x = r d_x$ = total number of grandprogeny (F_2) produced by the daughters of all n founding females (group productivity)
 π_x = probability that a particular group is founded by x A-type and $(n - x)$ B-type females

II Integration of individual selection with group selection

In the global population, the distribution of group composition is assumed to be binomial, so the probability that a group will have x A-type females among its n founders is

$$\pi_x = \binom{n}{x} \bar{p}^x (1 - \bar{p})^{n-x} \quad (5)$$

The average number of F_2 gene copies, per group, due to all founding females is

$$\bar{V} = \sum_{x=0}^n \pi_x V_x = n r^2 [\bar{p}(1 - a) + (1 - \bar{p})(1 - b)] \quad (6)$$

since $E[p_x] = \bar{p}$. The average number of F_2 gene copies, per group, due to A-type founding females is

$$\sum_{x=0}^n \pi_x p'_x V_x \quad (7)$$

Therefore, the proportion of grandmaternal gene copies in the F_2 that are due to A-type founding females in the global population is

$$\bar{p}' = \frac{\sum_{x=0}^n \pi_x p'_x V_x}{\bar{V}} \quad (8)$$

For example, for $n = 2$,

$$\bar{p}' = \frac{\bar{p}(1 - \bar{p})[(2a + b)(1 - a) + a(1 - b)] + \bar{p}^2(1 - a)}{\bar{p}(1 - a) + (1 - \bar{p})(1 - b)} \quad (9)$$

female-biased sex ratios for a population structured into randomly initiated groups, each founded by n inseminated females, whose offspring mate locally among themselves and then disperse to found new groups. For such a population, Hamilton showed that an autosomal gene that causes its female bearer to produce ar sons among her r progeny, where $a = (n - 1)/2n$, will be 'unbeatable' by a gene coding for any other sex ratio (diploidy, complete mating and equal investment per offspring assumed). Thus, the optimal progeny sex ratio is $1/4$ sons with $n = 2$ founding females, rising asymptotically toward $1/2$ sons with increasing n . A solitary founding female ($n = 1$) should produce the minimum number of sons required to inseminate her daughters.

Hamilton¹ and others^{4-6,8} have suggested that daughter-biased progenies are an adaptive response to selective pressures on the mother, serving to reduce 'local mate competition' (LMC) among her sons. By this argument, the increasing probability that brothers will compete for the same mates, as the number of sibships becomes smaller, is countered by a progressive decrease in the proportion of sons. Members of a local breeding group, such as the wasps in a single fig² or host pupa⁶, or the hummingbird flower mites in a single inflorescence³ are highly inbred, relative to the global population of which they are a part. Hamilton's¹ early emphasis on inbreeding, and Maynard Smith's⁴ reformulation of Hamilton's LMC model in terms of the proportion of matings that are between siblings, may have suggested to some authors⁸⁻¹⁰ that inbreeding (or sibmating) itself somehow provides a mechanism for the evolution of

daughter-biased progenies as an individually advantageous adaptation, reducing sibling competition for mates.

For the case of more than one founding female ($n > 1$), a simple algebraic model for the evolution of progeny sex ratio (Table 1) indicates that (1) LMC itself cannot select for female-biased progenies, (2) inbreeding (sibmating) is not relevant to the problem, and (3) group selection (differential productivity of genetically different groups) is required for the evolution of female-biased sex ratios.

In the model, each generation of local mating among the progeny of n founding mothers is followed by population-wide mixing and dispersal of their inseminated daughters. Two maternal phenotypes ('A-type' and 'B-type') compete, each producing a different, fixed progeny sex ratio (a and b , respectively, as proportion of sons). The fitness of a founder is measured in terms of the gene copies she leaves among grandprogeny^{1,4,6,11}. Figure 1 shows that the fitness of a 'fisherian' founder, producing an unbiased progeny sex ratio ($b = 0.5$), always exceeds the fitness of a 'hamiltonian' founder, producing a daughter-biased sex ratio ($a < 0.5$), regardless of the number of founders, initial proportion of the two types and degree of female bias in hamiltonian progenies¹. Because the degree of selective advantage of fisherian females depends only on their proportion among founders, relative fitness of the two types is completely unaffected by the level of inbreeding or sibmating, and by the intensity of competition among brothers for mates, which are functions of the number of founders, given a and b . Individual selection on females whose offspring breed in local

groups with the offspring of other founders, however few, cannot favour daughter-biased progenies.

How, then, can female-biased sex ratios evolve? If the dispersal pool includes more than one sex-ratio phenotype, local groups will vary by chance in the frequency of the phenotypes among founders. Within each group, the relative fitness of fisherian founders exceeds that of hamiltonian founders. However, groups with higher-than-average initial frequencies of hamiltonian females produce a greater-than-average absolute number of grandprogeny, simply because their founders jointly produce more daughters (and fewer sons) than

Table 2 Individual and group selection in the evolution of female-biased sex ratios: a numerical example with $n = 4$

	Group 1	Group 2
I Initial composition of groups		
A. Number of hamiltonian ($a = \frac{1}{2}$) founders: x	3	1
B. Number of fisherian ($b = \frac{1}{2}$) founders: $(n - x)$	1	3
C. Proportion hamiltonian founders: p_x	0.75	0.25
II Contribution of founders to grandprogeny:		
Individual selection		
A. Proportion of grandmaternal gene copies due to		
1. Each hamiltonian founder	0.2470	0.2471
2. Each fisherian founder	0.2591	0.2510
3. All hamiltonian founders: p'_x	0.7409	0.2471
4. All fisherian founders: $(1 - p'_x)$	0.2591	0.7529
B. Fitness of a hamiltonian founder relative to a fisherian founder	0.9533	0.9845
III Differential productivity of groups:		
Group selection		
A. Total grandprogeny produced by each group: V_x	$2.375r^2$	$2.125r^2$
B. Number of grandprogeny gene copies within each group due to		
1. Each hamiltonian founder	$0.5865r^2$	$0.5251r^2$
2. Each fisherian founder	$0.6154r^2$	$0.5333r^2$
3. All hamiltonian founders: $p'_x V_x$	$1.7596r^2$	$0.5251r^2$
4. All fisherian founders: $(1 - p'_x) V_x$	$0.6154r^2$	$1.5999r^2$
IV Number of gene copies among pooled grandprogeny		
A. Due to each hamiltonian founder: $\frac{1}{4}(1.7596 + 0.5251)r^2 =$	$0.5712r^2$	
B. Due to each fisherian founder: $\frac{1}{4}(0.6154 + 1.5999)r^2 =$	$0.5538r^2$	

Individual and group selection in the evolution of female-biased sex ratios. This example is based on an arbitrary initial composition of groups. Two groups are formed by sampling from a founder pool that consists of half fisherian (F) and half hamiltonian (H) inseminated females. The groups differ in the frequency of the two types (line IC) but together reflect the (equal) frequencies in the founder pool. Each female in each group then produces r progeny, of which $1/2 r$ are sons, for each fisherian mother, and $3/8 r$ are sons, for each hamiltonian mother. The progeny mate at random among themselves within each group (all females are inseminated). Computing fitness within groups as the expected number of grandmaternal gene copies due to each founding female (lines IIB, 1, 2), it is evident that the individual fitness of each fisherian female exceeds that of each hamiltonian female within groups (line IIB). This is true for any mixed group (Fig. 1). On the other hand, group fitness—the total number of grandprogeny produced (line IIIA)—is greater for the group with the higher initial frequency of hamiltonian females, because its founders jointly produce more daughters (and fewer sons) than do the predominantly fisherian founders of the other group. Because the more productive group is mostly hamiltonian, when the number of gene copies due to each of the two types of founders is evaluated among the pooled grandprogeny of the two groups, the hamiltonian founders prove to have contributed more than the fisherian founders (lines IVA, B), in spite of the selective disadvantage of hamiltonian founders within each group. The force of group selection depends on the variance in initial composition of groups, here arbitrarily established to make a point. In fact, random (binomial) variation is sufficient for the force of group selection to exceed that of individual selection^{3,13}. Greater-than-binomial variance only increases the effect^{2,13}. In terms of Price's^{2,14} hierarchical model of selection, the fitness of groups (V_x) co-varies positively with the proportion of hamiltonian founders that initiate them (p_x), whereas the fitness of individuals within groups co-varies negatively with possession of the hamiltonian trait (or, genetically, with the number of hamiltonian alleles an individual carries): p'_x is inevitably less than p_x . Thus, female-biased sex ratios are opposed by individual selection, but are favoured by group selection.

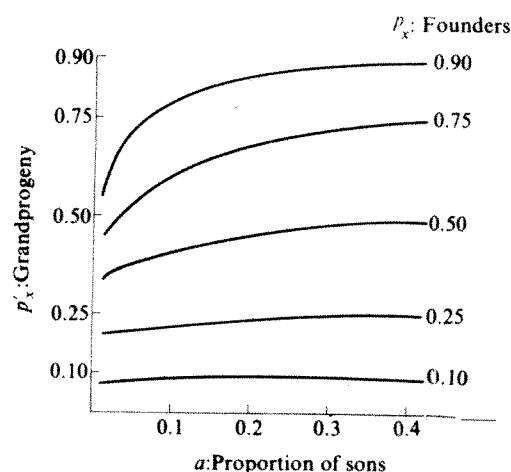


Fig. 1 Fisher's⁷ sex ratio principle. The selective disadvantage of producing female-biased progeny. A group is founded by hamiltonian females in proportion p_x , each producing proportion $a < 0.5$ sons, and by fisherian females in proportion $(1 - p_x)$, each producing half sons and half daughters. Here, the proportion of grandmaternal gene copies in the grandprogeny that are due to the hamiltonian founders, p'_x (computed by equation (3), Table 1) is plotted as a function of a for several values of p_x . The figure demonstrates that p'_x is inevitably less than p_x : the fitness of fisherian founders always exceeds that of hamiltonian founders in a mixed group. Note that the number of founding females (n , group size) plays no part in this result. Fisher's principle operates with equal force in random-mating groups or populations of any size, whether initiated

average. The crucial point is this: when genetic contributions to the grandprogeny of randomly initiated groups are pooled by dispersal, hamiltonian founders prove to have made a greater per capita contribution than fisherian founders, even though within every mixed group the reverse is true. Table 2 provides a numerical example of this phenomenon.

In general terms, the global proportion of grandmaternal gene copies in the grandprogeny that are due to A-type founders is a productivity-weighted average of local (within-group) proportions, derived in part II of Table 1 on the assumption of random (binomial) variation among groups in the initial frequency of A-type and B-type founders. When A-type females are optimally hamiltonian [$a = (n - 1)/2n$] it is easily shown numerically (for example, with equation (9), Table 1) that the global proportion of grandmaternal gene copies due to hamiltonian females is greater than the proportion of hamiltonian females among group founders, regardless of the progeny sex ratio produced by B-type founders and initial frequencies of the two types. (Hamilton¹ obtained this result analytically by direct optimization.) In other words, whatever the mode of inheritance, optimally hamiltonian alleles will increase in frequency relative to fisherian (or any other) alleles in a randomly structured population. The present model permits a more general statement: if $b < a \leq (n - 1)/2n$ or if $b > a \geq (n - 1)/2n$, A-type females prevail at any starting frequency; if $a < (n - 1)/2n < b$, a frequency-dependent stable polymorphism exists (for example, if $n = 3$, $a = 0.1$, and $b = 0.6$, then $\bar{p}' = \bar{p} = 0.48$).

The force of individual selection is independent of the number of founders (n) (Fig. 1), whereas the force of group selection increases with genetic variation among groups^{12,13}, which becomes greater as n decreases, producing increasingly female-biased sex ratios. If among-group variation in progeny sex ratio is suppressed, or if it is permitted, but differential productivity is eliminated by the imposition of a carrying capacity, female-biased sex ratios cannot evolve, regardless of n , and therefore regardless of the level of inbreeding, sibmating and local mate competition³.

I thank D. S. Wilson for consultations and encouragement, and H. J. Brockmann, E. L. Charnov, A. B. Clark, L. Goldwasser, M.-C. King, S. Levin, J. Maynard Smith, P. W. Sherman, N. C. Stenseth and J. H. Werren for stimulating discussions. This work was supported by the NSF (DEB78-12038).

Received 4 August 1980; accepted 19 January 1981.

1. Hamilton, W. D. *Science* **156**, 477-488 (1967).
2. Hamilton, W. D. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. S. & Blum, N. A.) 167-221 (Academic, New York, 1979).
3. Wilson, D. S. & Colwell, R. K. *Evolution* (in the press).
4. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
5. Bulmer, M. G. & Taylor, P. D. *Nature* **284**, 448-449 (1980); *J. theor. Biol.* **86**, 83-89 (1980).
6. Wernick, J. H. *Science* **208**, 1157-1159 (1980).
7. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
8. Alexander, R. D. & Sherman, P. W. *Science* **196**, 494-500 (1977).
9. Trivers, R. L. & Hare, H. *Science* **191**, 249-263 (1976).
10. Stenseth, N. C. *Oikos* **30**, 83-89 (1978).
11. Shaw, R. F. & Mohler, J. D. *Am. Nat.* **87**, 337-342 (1953).
12. Slatkin, M. & Wade, M. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3531-3534 (1978).
13. Wilson, D. S. *The Natural Selection of Populations and Communities* (Benjamin/Cummings, Menlo Park, 1980).
14. Price, G. A. *Ann. hum. Genet.* **35**, 485-490 (1972).

Growth hormone stimulates adrenal steroidogenesis in the fetus

Uday P. Devaskar*, Sherin U. Devaskar,
Sandra Voina, Nelson Velayo & Mark A. Sperling

Department of Pediatrics, Division of Neonatology and
Endocrinology, Children's Hospital Research Foundation,
University of Cincinnati Medical Center, Cincinnati,
Ohio 45267, USA

The 'surge' of corticosteroid in fetal plasma during late gestation has been implicated in the initiation of parturition and the maturation of enzyme systems in organs such as the lung, liver, adrenal medulla and thyroid¹⁻⁵. But in all species studied, the mechanism responsible for increased secretion in the fetus remains unclear. The hormone adrenocorticotropin (ACTH) is well established as the primary regulator of adrenocortical cellular growth and secretory function during fetal and adult life⁶. However, no increase in fetal plasma ACTH has been observed before the corticosteroid surge in sheep or humans^{7,8}, although differential responsiveness of the fetal adrenal cortex to ACTH at various gestational ages and a possible role of extra-adrenal inhibitory factors have been proposed⁹⁻¹². The possible steroidogenic effect of α -melanocyte-stimulating hormone (α MSH) on fetal adrenocortical function is controversial¹³⁻¹⁶, and we could not demonstrate any such action in fetal lamb or rabbit^{17,18}. Prolactin and growth hormone (GH) potentiate the steroidogenic effect of ACTH in adult rats¹⁹⁻²² but prolactin had no such effect in fetal lamb¹⁷. We have investigated the steroidogenic properties of GH, and report here that it stimulates adrenal steroidogenesis in the fetal but not maternal rabbit *in vitro* and in fetal but not maternal sheep *in vitro* as well as *in vivo*.

We used pregnant New Zealand White rabbits of known gestational age (g.a. 25-30 days, term 31 days, $n=8$) and Columbia-Suffolk strain ewes (g.a. 140-144 days, term 150 days, $n=5$). Each was killed by intravenous pentobarbitone and after hysterotomy, the fetuses were rapidly removed. Adrenal glands were removed *in toto*, without disrupting the capsule, and bathed in culture medium M199 with Hank's salt solution (Gibco) containing 2.2 mM calcium chloride. Adrenals were pooled within each litter of rabbits. Glands were trimmed of fat and connective tissue, minced when necessary and bathed in 5 ml of medium M199 containing 2.5 mg ml⁻¹ of lyophilized trypsin (Worthington), which digests adrenal medullary cells but leaves adrenocortical cells intact. We previously described¹⁸ the details of isolating dispersed viable adrenocortical cells by a modification of Sayers' technique²³. We then incubated 3×10^4 cells in 150 μ l (rabbit) or 10^5 cells in 200 μ l (sheep) of the culture

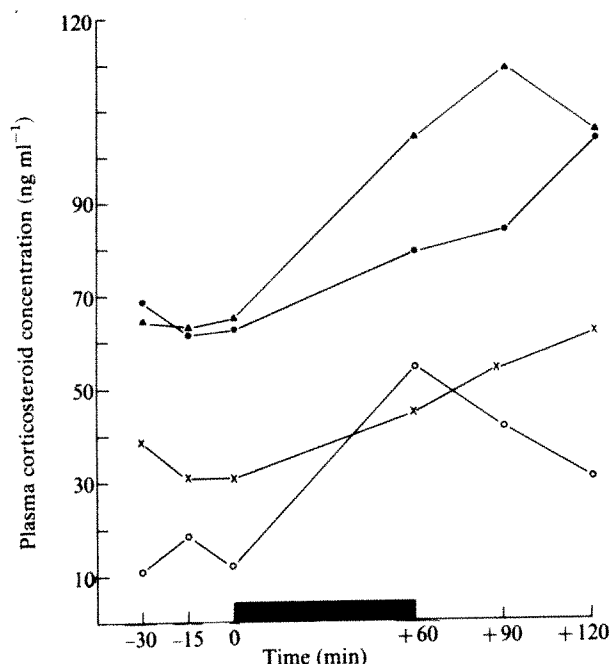


Fig. 1 Fetal lamb plasma corticosteroid concentrations before and after HGH infusion (shown by shaded bar). Age of lambs: ▲, 138 days; ●, 143 days; ×, 139 days; ○, 137 days.

medium, in duplicate with 50 μ l of phosphate buffer (0.01 M monobasic sodium phosphate, 0.9% NaCl, pH 7.4), adding human GH (HGH, National Pituitary Agency, NIH-HS 2243E) or synthetic ACTH₁₋₂₄ (Organon), the biologically active portion of the intact hormone ACTH₁₋₃₉, in the buffer. The incubation was performed in polystyrene tubes (12 $\times$ 75 mm) in a water bath at 37 °C under a mixture of 5% CO₂ and 95% O₂ for 2 h, after which the cells were removed by centrifugation at 300g for 10 min. Aliquots (50 μ l) of supernatant were removed, in duplicate, to measure corticoid production. Response to ACTH₁₋₂₄ served as the control for the steroidogenic response of the cell preparation. Corticosteroids were measured by a specific radioimmunoassay (RIA)^{24,25}. Because rabbit adrenals secrete both cortisol and corticosterone²⁶, an antibody (Radioassay System Catalogue 147) that reacts 100% with both hormones was used with rabbit adrenal cells, but because cortisol is the major corticoid in sheep²⁷, an antibody (Radioassay System Catalogue 137) that cross-reacts 100% with cortisol and minimally with other corticoids was used with sheep adrenal cells. The results represent total corticosteroid production by dispersed rabbit and sheep adrenocortical cells.

For *in vivo* studies, pregnant Columbia-Suffolk strain ewes of known gestational age and adult non-pregnant sheep were used. Pregnancy was confirmed by X-ray analysis between 85 and 90 days of gestation. The maternal and fetal carotid artery and jugular vein were catheterized²⁸ at least 1 week before the experiment. At 137, 138, 139 and 143 days of gestation, 20 μ g

Table 1 Basal corticosteroid production and increment above basal from hormone stimulation by fetal and maternal rabbit adrenocortical cells

Corticosteroid production (ng per 3×10^4 cells per 2 h)		
	Fetus ($n=8$)	Mother ($n=4$)
Basal	3.69 ± 0.66	3.38 ± 0.52
Increment above basal		
ACTH (25 μ g ml ⁻¹)	2.69 ± 0.32	2.80 ± 0.26
ACTH (125 μ g ml ⁻¹)	3.96 ± 0.47	3.85 ± 0.17
HGH (1.25 μ g ml ⁻¹)	3.30 ± 0.37	-0.05 ± 0.08
HGH (2.5 μ g ml ⁻¹)	4.59 ± 0.27	-0.04 ± 0.05

Values are mean $\pm$ s.e.m.

* To whom reprint requests should be addressed.

of HGH in 8 ml of normal saline was infused for 1 h through the fetal jugular vein. Fetal arterial blood (1 ml) was drawn at timed intervals before and after starting HGH infusion as indicated in Fig. 1; pH, pO_2 and pCO_2 were measured before and after infusion and were similar and in the normal range of our experimental animals. To exclude the effects of stress due to saline infusion, two animals earlier infused with HGH were infused with 8 ml of normal saline through the fetal jugular vein at 140 and 141 days of gestation and fetal corticosteroid concentration in plasma measured. We also infused HGH into catheterized adult sheep ($n = 3$) using the same protocol. With-drawn blood was immediately centrifuged at 500g for 10 min at 4°C and plasma stored at -20°C until measurement of corticosteroid concentrations²⁵. Plasma concentrations of HGH achieved *in vivo* were measured by RIA²⁹ in two fetal and one adult sheep before and after HGH infusion.

The mean basal production of corticoids, and the increase (mean $\pm$ s.e.m.) due to hormonal stimulation in rabbit and sheep adrenal cells are shown respectively in Tables 1 and 2. Although fetal cells from both species significantly increased steroidogenesis under the influence of ACTH₁₋₂₄ or HGH, the maternal cells responded to ACTH₁₋₂₄ only. Furthermore corticoid production by fetal rabbit and fetal sheep cells after incubation with 2.5 μ g ml⁻¹ of HGH was significantly higher than after 1.25 μ g ml⁻¹ (paired *t*-test $P < 0.02$).

Basal plasma corticosteroid concentrations in fetal lamb were significantly higher than in adult sheep (Figs 1, 2), but in both were within range reported by others³⁰⁻³². Plasma corticosteroid concentrations increased between 60 and 350% above basal after infusion of HGH in all four fetal lambs (Fig. 1). No such increase was noted in fetuses infused with saline or in adults infused with HGH (Fig. 2). Basal fetal and adult plasma HGH concentrations were less than the sensitivity of the assay (0.1 ng ml⁻¹), indicating minimal cross-reaction between endogenous ovine growth hormone and the HGH, especially as GH concentrations in ovine fetal plasma are high^{33,34}. Fetal lamb plasma HGH concentrations at the end of infusion (+60 min) were 15.3 and 25 ng ml⁻¹ while adult sheep plasma HGH concentration was 2.1 ng ml⁻¹.

The HGH we used is a highly purified immunochemical. Because maternal adrenal cells showed no steroidogenic response, unlike fetal cells, contamination of HGH with biologically active ACTH seemed unlikely. To confirm this assumption, the buffer and 2.5 μ g of HGH in the buffer were assayed for ACTH by RIA at Nichols' Institute³⁵. The ACTH concentrations in HGH plus buffer were the same as those in the buffer alone. Contamination with α MSH seemed unlikely for the following reasons. Using the same dispersed cell system and RIA techniques, we have observed no steroidogenic effect of α MSH in concentrations from 8 to 800,000 pg ml⁻¹ in either fetal sheep or rabbits at various gestational ages^{17,18}. Similarly, infusion of 75 μ g of synthetic α MSH (Bachem) in chronically catheterized fetal sheep at 127-130 days of gestation ($n = 4$) produced no increase above basal fetal plasma corticosteroid concentrations¹⁷. As a direct steroidogenic effect of HGH was observed *in vitro*, the increase in fetal plasma corticosteroid

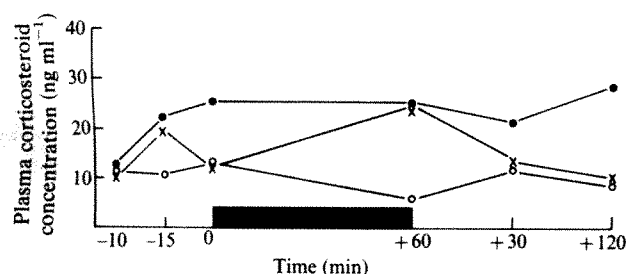


Fig. 2 Adult sheep plasma corticosteroid concentrations before and after HGH infusion (shown by shaded bar).

Table 2 Basal corticosteroid production and increment above basal from hormone stimulation by fetal and maternal sheep adrenocortical cells

Corticosteroid production (ng per 10 ⁵ cells per 2 h)		
Basal	Fetus ($n = 5$)	Mother ($n = 3$)
Increment above basal	1.79 $\pm$ 0.42	1.92 $\pm$ 0.59
ACTH (25 pg ml ⁻¹)	2.87 $\pm$ 0.63	5.04 $\pm$ 1.54
HGH (1.25 μ g ml ⁻¹)	3.08 $\pm$ 0.86	-0.15 $\pm$ 0.17
HGH (2.5 μ g ml ⁻¹)	5.33 $\pm$ 1.1	0.13 $\pm$ 0.19

Values are mean $\pm$ s.e.m.

observed *in vivo* is unlikely to have been due to ACTH released from the fetal pituitary.

Although no data are available for the rabbit, fetal plasma GH concentrations are very high compared with those during neonatal and adult life in sheep^{33,34} and humans³⁶. Because GH does not cross the placenta and the metabolic clearance rates in the fetal and neonatal period are similar³⁷, high fetal plasma concentrations probably represent increased secretion by the fetal adenohypophysis. The physiological significance of high fetal plasma GH concentrations is unclear; it does not seem to be essential for somatic growth³⁷. Our findings provide the first experimental evidence that GH directly influences adrenal steroidogenesis during fetal life. Although concentrations of HGH achieved *in vivo* were within the physiological concentrations of ovine GH observed in fetal (up to 250 ng ml⁻¹) and adult sheep plasma^{33,34}, the concentrations of HGH tested *in vitro* were one log order greater. Using entire adrenal glands rather than dispersed cells, collected from fetal rabbits at 24-27 days of gestation, Rudman *et al.* observed no increase in corticosterone production due to HGH stimulation¹⁶. The discrepancy with our results may be due to the different design of the experiments, the different gestation, the different RIA or a combination of these factors.

Received 9 June 1980; accepted 21 January 1981.

- Liggins, G. C. *Am. J. Obstet. Gynec.* **126**, 931 (1976).
- Thornburn, G. D. & Challis, J. R. G. *Physiol. Rev.* **59**, 863 (1979).
- Nathanielsz, P. W. A. *Rev. Physiol.* **40**, 411 (1978).
- Ballard, P. L. *Monographs on Endocrinology: Glucocorticoid Hormone Action* (eds Baxter, J. D. & Rousseau, G. G.) 493 (Springer, Berlin, 1979).
- Thomas, A. L., Crane, E. J. & Nathanielsz, P. W. *Endocrinology* **103**, 17 (1978).
- Sayers, G. & Portanova, R. *Handbook of Physiology* Vol. 5 Sect. F (eds Green, R. D. & Astwood, E. B.) 41 (Am. Physiol. Soc., Washington DC, 1975).
- Rees, L., Jack, P. M. B., Thomas, A. L. & Nathanielsz, P. W. *Nature* **253**, 274 (1975).
- Winters, A. J., Oliver, C., Colston, C., MacDonald, P. C. & Porter, J. C. *J. clin. Endocr. Metab.* **39**, 269 (1976).
- Glickman, J. A. & Challis, J. R. G. *Endocrinology* **106**, 1371 (1980).
- Madill, D. & Bassett, J. M. *J. Endocr.* **58**, 75 (1973).
- Roebuck, M. M., Jones, C. T., Holland, D. & Silman, R. *Nature* **284**, 616 (1980).
- Nathanielsz, P. W. *Fetal Endocrinology: An Experimental Approach*, 191 (North-Holland, Amsterdam, 1976).
- Llanos, A. J. *et al. Endocrinology* **105**, 613 (1979).
- Walsh, S. W., Norman, R. L. & Novy, M. J. *Endocrinology* **104**, 1805 (1979).
- Jujeda, K., Faiman, C., Reyes, F. I., Thliveris, J. & Winter, J. S. D. *Abstr. Endocrine Soc. Mtg* **626**, 229 (1979).
- Rudman, D., Hollins, B. M., Lewis, N. C. & Chawla, R. K. *J. clin. Invest.* **65**, 822 (1980).
- Magyar, D. M., Devaskar, U. P., Fridshal, D., Buster, J. E. & Nathanielsz, P. W. *Endocrinology* **107**, 1582 (1980).
- Devaskar, U. P., Magyar, D., Fridshal, D., Buster, J. E. & Nathanielsz, P. W. *Endocrinology* **107**, 809 (1980).
- Lis, M., Gilardeau, C. & Cheretien, M. *Clin. Res.* **21**, 1027 (1973).
- Colby, H. D., Caffrey, J. L. & Kitay, J. I. *Endocrinology* **93**, 188 (1973).
- Kramer, R. E., Crier, J. W. & Colby, M. D. *Endocrinology* **101**, 297 (1977).
- Coyne, M. D. *Abstr. Endocr. Soc. Mtg* **632**, 230 (1979).
- Sayers, G., Swallow, R. L. & Giordano, N. D. *Endocrinology* **88**, 1063 (1971).
- Radioassay Systems Laboratories, Inc. *Direct Corticosterone by RIA* (1979).
- Radioassay Systems Laboratories *Cortisol ¹²⁵I Kit Cat.* **137**, 133 (1978).
- Mulay, S., Giannopoulos, G. & Solomon, S. *Endocrinology* **93**, 1342 (1973).
- Thomas, S. J., Wilson, D. W., Pierrepont, C. G., Cameron, E. H. D. & Griffiths, K. T. J. *Endocr.* **68**, 181 (1976).
- Sperling, M. A., Christensen, R. A., Ganguli, S. & Anand, R. *Pediatr. Res.* **14**, 203 (1980).
- Goodman, H. G., Grumbach, M. M. & Kaplan, S. L. *New Engl. J. Med.* **278**, 56 (1968).
- Strott, C. A., Sundel, H. & Stahlman, M. T. *Endocrinology* **95**, 1327 (1974).
- Brown, E. H., Coghlan, J. P., Hardy, K. J. & Wintour, E. M. *Acta endocr. Copenh.* **88**, 364 (1978).
- Magyar, D. M. *et al. Endocrinology* **107**, 155 (1980).
- Bassett, J. M., Thornburn, G. D. & Wallace, A. L. *J. Endocr.* **48**, 251 (1970).
- Gluckman, P. D., Mueller, P. L., Kaplan, S. L., Rudolph, A. M. & Grumbach, M. M. *Endocrinology* **104**, 162 (1979).
- Nichols, A. L. & Nelson, J. C. *Radioimmunoassay Manual* Vol. 1 (Nichols' Institute, 1977).
- Kaplan, S. L., Grumbach, M. M. & Shephard, T. H. *J. clin. Invest.* **51**, 2080 (1972).
- Nathanielsz, P. W. *Fetal Endocrinology: An Experimental Approach*, 108 (North-Holland, Amsterdam, 1976).

The respiratory burst of phagocytic cells is associated with a rise in vacuolar pH

Anthony W. Segal*, Michael Geisow†, Rodolfo Garcia*, Angela Harper* & Robert Miller‡

* Department of Haematology, School of Medicine, and
† Department of Biophysics, NIMR, Mill Hill, London NW7 1AA, UK
‡ Department of Zoology, University College London, London WC1E 6BT, UK

Neutrophil leukocytes are the body's major defence against bacteria, which they phagocytose and kill. It has been found that phagocytosis and killing are accompanied by a dramatic rise in non-mitochondrial respiration^{1,2}; and that the efficiency of killing is impaired in the absence of oxygen^{3,4}. It is also impaired in neutrophils from patients with chronic granulomatous disease (CGD), where the respiratory burst is absent⁵. This has been difficult to reconcile with their normal content of granule proteins that kill bacteria *in vitro*⁶. Indeed, CGD cells are essentially normal both morphologically and constitutionally^{7,8} except that they lack a functional very low potential cytochrome *b* (*b*₂₄₅)⁹, which is a component of the oxidase system responsible for the respiratory burst of normal cells^{10,11}. Activation of the oxidase is associated with the generation of various reduced oxygen species¹²⁻¹⁶ which have been widely thought to be responsible for the killing of phagocytosed microorganisms either directly¹⁷, or by acting as substrates for myeloperoxidase-mediated halogenation¹⁸. We report here, however, that a major consequence of the defective function of this oxidase in neutrophils and monocytes from CGD patients is an absence of the normal initial rise, and an unusually rapid and extensive fall in pH which is itself associated with the impairment of the killing and digestion of intracellular staphylococci.

Phagocytosis and bacterial killing were measured in cells from five severely affected male patients with X-linked CGD, all of whose cells lacked spectral evidence of cytochrome *b*₂₄₅ (ref. 9), a mildly affected girl with the autosomal recessive variant⁹ and normal subjects. Phagocytosis was assayed by lysostaphin lysis of extracellular organisms¹⁹. It was very rapid in normal cells (*t*_{1/2} = 0.65 min) and maximal at 4 min (Fig. 1). The bacteria then re-emerged from the cells as shown by an increase in extracellular radioactivity from 13% at 4 min to 32% at 16 min. This expulsion of phagocytosed material, mostly in the form of particles which sedimented after gentle centrifugation, has been described previously²⁰. Phagocytosis by the patients' cells was slightly slower (*t*_{1/2} = 1.26 min), and there was no evidence of the subsequent release of bacteria from the cells (Fig. 1).

Bacterial killing by normal cells was very rapid (Fig. 1). The patients' cells also killed more than half the organisms and when corrections were made for extracellular organisms it was calculated that no bacteria remained viable within normal cells and that only 22% of those within the patients' cells survived.

To what can one attribute the reduced efficiency of bacterial killing by CGD cells? We have previously proposed that one role of the vacuolar cytochrome system is to regulate the pH within phagocytic vacuoles^{11,21}, and there is evidence that the pH falls faster after phagocytosis in CGD than in normal cells²². Thus, we measured early pH changes in the phagocytic vacuoles of leukocytes from healthy and granulomatous individuals, using the pH indicator fluorescein²³ conjugated to *Staphylococcus aureus*. Glass coverslips were coated with monolayers (1×10^4 cells mm⁻²) of neutrophils (>95% pure) and mononuclear cells (>80% monocytes) which were then exposed to a suspension of fluoresceinated bacteria at 4 °C. The pH of the bacterial environment was determined by measuring fluorescein emission

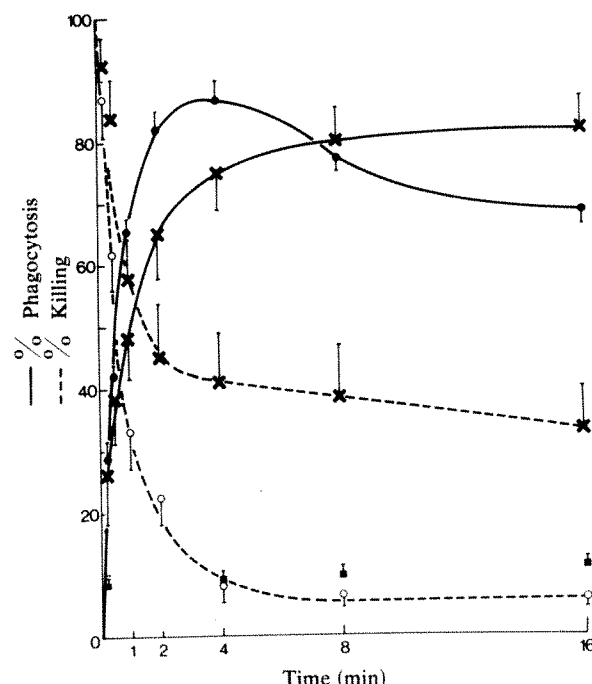


Fig. 1 Phagocytosis and killing of staphylococci by normal (●, ○) and CGD (×) neutrophils. Solid line, intracellular bacteria; broken line, viable bacteria (%). *S. aureus* (Oxford Strain NCTC 6571) from an overnight culture in broth (Oxoid) containing 25 μ Ci ml⁻¹ ³H-glucose (Radiochemical Centre, TRK. 85) were opsonized with human IgG (ref. 41) and added (in a 1:1 ratio) to 1×10^8 purified neutrophils (>95% in normals, patients' cells containing 5–10% eosinophils) in 1.5 ml RPMI 1640 medium (Flow Labs) containing 5 IU ml⁻¹ heparin in a rapidly stirred thermostatically controlled chamber⁴¹. Aliquots (50 μ l) were taken into 500 μ l ice-cold RPMI containing *N*-ethyl maleimide (1 mM) and 100 μ g ml⁻¹ lysostaphin (Sigma). The samples were mixed, incubated at 37 °C for 15 min, recooled and centrifuged at 7,000g for 4 min. Radioactivity in the supernatant was measured in a liquid scintillation counter. Fifty-microlitre aliquots of the mixture of bacteria and cells were taken at the same times into 10 ml ice cold H₂O. After 1 h aliquots were plated out for colony counts. Results are expressed as the mean $\pm$ s.e. of seven studies on cells from six CGD patients, each study being controlled with a simultaneous study on normal cells. Lactate dehydrogenase (mean $\pm$ s.e. of four experiments) release³⁰ from normal cells is shown (■).

at selected excitation wavelengths at various times after the initiation of phagocytosis. In normal neutrophils an increase in peak fluorescence corresponding to an apparent rise of intravacuolar pH to 7.75 was observed within the first 2 min (Fig. 2). Thereafter the pH fell slowly to 6.0–6.5 after 2 h. The same pattern was observed with normal monocytes (Fig. 2b). Similar studies with yeasts impregnated with the pH indicator bromothymol blue^{22,24} also showed that almost immediately following phagocytosis by normal cells there was a transient colour change from yellow to blue indicating a rise in pH from 6.8 in the medium to >7.6.

The pattern of pH changes with phagocytosis was clearly different in CGD leukocytes. There was neither colour change with bromothymol blue-impregnated yeasts, nor initial fluorescence increase with fluorescein-labelled *S. aureus*; instead the pH fell rapidly from 7.4 to 6.7 within 2 min. The final vacuolar pH of 5.5 was lower than that of normal cells. Addition of methylamine, which elevates the pH of secondary lysosomes²³, to the cell perfusate and the pre-incubation medium, effectively prevented this early rapid drop in vacuolar pH (Fig. 2), although phagocytosis was not inhibited.

Furthermore, spectra from fluorescent *S. aureus* bound to CGD or normal neutrophils at 10 °C (before phagocytosis) were identical to the spectrum of the free bacterium at pH 7.4, whereas after 13 min incubation at 37 °C the spectra were

completely different; that of CGD cells could be fitted best by the pH 5.9 spectrum of fluorescein and that of normal cells by the pH 7.0 spectrum (data not shown).

When the experiment was repeated with normal human neutrophils in anaerobic conditions, achieved by perfusing the cells with degassed buffers or mixing a concentrated suspension of neutrophils with the fluorescent bacteria at 4 °C in fine capil-

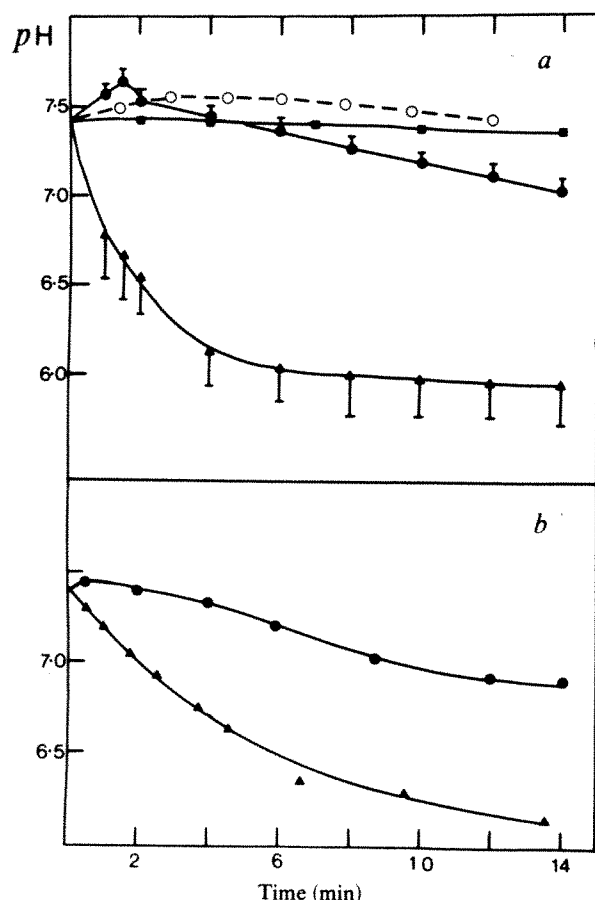


Fig. 2 Phagosomal pH in leukocytes after phagocytosis. *a*, Neutrophils: mean $\pm$ s.d. of two studies on each of two patients with CGD and one study on a third ($\blacktriangle$) and on five paired control subjects ($\bullet$). The effects of methylamine²³ (10 mM) on the pH within CGD cells ($\blacksquare$) and cytochalasin β ($1 \mu\text{g ml}^{-1}$) on control cells ($\circ$) are also shown. *b*, Monocytes: mean results of studies on two CGD patients ($\blacktriangle$) and paired control subjects ($\bullet$). *S. aureus* (50 μl of packed organisms) was well washed with saline and 50 mM pyridine-HCl buffer, pH 5.0. The bacteria were resuspended in 1 ml of the same buffer and mixed with 40 μl of fluorescein amine isomer I (Sigma), 1 mg ml^{-1} in dioxan. One mg of 1-ethyl 3-diaminopropyl carbodiimide was added and the suspension agitated for 2 h. The suspension was washed first in the pyridine buffer containing 100 mM glycine and then in saline. Fluorescence microscopy showed that the killed bacteria had been uniformly surface labelled. Freshly isolated human neutrophils or mononuclear cells⁴² were allowed to attach to acid-etched glass coverslips for 15 min at room temperature. Non-adherent cells were then washed away with RPMI medium, and the resulting monolayers were stored under this medium. Fluorescent *S. aureus* opsonized with IgG (ref. 41) were allowed to sediment on to leukocyte monolayers and attach at 4 °C for 20 min. Free or loosely bound labelled bacteria were removed by vigorous washing in ice-cold saline or RPMI medium. Fluorescence monitoring with a Perkin-Elmer MPF 4 fluorimeter began immediately after transfer of the cold monolayer to pre-warmed fluorescence cuvettes perfused with oxygenated RPMI medium at 37 °C. Phagosomal pH was calculated from the shape of the pH-dependent excitation spectrum of fluorescein by comparison with calibration spectra over a pH range of 4–8 from the same fluorescent bacteria adsorbed to glass in the absence of cells. The excitation spectra from fluorescent bacteria were corrected for the intrinsic fluorescence of the leukocytes and for scattered light.

larities which were then sealed, the initial pH rise was abolished and the pH fell rapidly, just as in the CGD patients' cells.

In these experiments bacteria were layered on the monolayers of cells at 4 °C to permit attachment without phagocytosis, which was then synchronized by rapid warming to 37 °C. To ensure that cooling the cells to 4 °C was not modifying the results, the cells were incubated with bacteria at 25 °C for 30 s before vigorous washing and pH measurement. The same pH changes as described above were observed in normal and CGD cells. Pre-incubation with cytochalasin B (5 μM) retarded phagocytosis and diminished and delayed the pH changes (Fig. 2). Neither the direct action of the products of the oxidase, generated at pH 7.4 with a xanthine oxidase-hypoxanthine system⁹, nor the method of coupling the fluorescein to the organism, on carboxyl groups with carbodiimide or on amino groups with isothiocyanate, altered the fluorescence characteristics of fluorescein.

The fluorescence measurements include signals from both phagocytosed bacteria and extracellular organisms which remain in the buffered medium. As the constant signal from the latter minimizes the average change, and consequently that attributable to intracellular particles, it was important to measure the phagocytosis of bacteria by the monolayer cells, which was done by the release of radioactivity from extracellular bacteria after digestion with lysostaphin as described in Fig. 1 legend. Fifty per cent were taken up by 30 s, 80% by 2 min and a maximum of 10% finally remained extracellular. The kinetics of phagocytosis were the same in cells from a patient with CGD. In studies with fluorescent *S. aureus* the addition of lysostaphin after 15 min caused a small decrease in fluorescence, but insufficient to alter the pH profiles, confirming that the pH measurement pertained to organisms within the cell.

To determine whether the excessively acid conditions within the vacuole of CGD cells could explain the defective bacterial killing, we examined the killing of staphylococci by the cells of two patients in the presence and absence of 10 mM methylamine, which we had shown to elevate the pH in the vacuoles of patients' cells (Fig. 2). Regression analysis of the relationship between extracellular and viable organisms (see Fig. 1 legend) indicated that in the presence of methylamine (which did not itself kill bacteria) all the intracellular organisms were killed and that bacterial killing was significantly better ($P < 0.02$)²⁵ than in its absence.

To investigate other possible effects of the excessive acidity in CGD vacuoles we measured the digestion of radiolabelled, heat-killed staphylococci. Solubilization of the phenylalanine label was less than half of normal, that of the incorporated glucose was about 80% of normal whereas that of uridine label was normal (Fig. 3). Phagocytosis of the labelled bacteria by the two groups of cells was similar. The differences in the solubilization of bacteria seemed unrelated to differences in their content of hydrolytic enzymes because incubation of labelled bacteria with homogenates of normal or CGD cells at pHs of 5–8 released similar amounts of radiolabelled phenylalanine and glucose. Nor could the reduced solubilization in CGD cells be accounted for by less efficient degranulation. Phagocytic vacuoles, isolated by flotation²⁶, from four patients were in fact found to contain more enzymes than normal, the specific activities of β glucuronidase²⁷, lysozyme²⁸ and myeloperoxidase²⁹ being 4.55, 2.86 and 1.19 of those in vacuoles from the paired control cells. Half-maximal degranulation times for these enzymes, determined over 1 h, were normal²⁶.

We next investigated the difference between the expulsion of phagocytosed bacteria by normal and CGD cells (Fig. 1). We found that normal cells released three times more solubilized ³H-glucose label than cells of the four patients (Fig. 3a, $P < 0.0025$)²⁵, although cell damage as measured by lactate dehydrogenase release³⁰ after incubation for 2 h was not significantly different (5.7 ± 1.2 (s.e.) in controls and $8.0 \pm 3.9\%$ in patients). For morphological studies cells and bacteria were incubated in the conditions used in the killing assay (see Fig. 1 legend) and timed samples were collected for electron

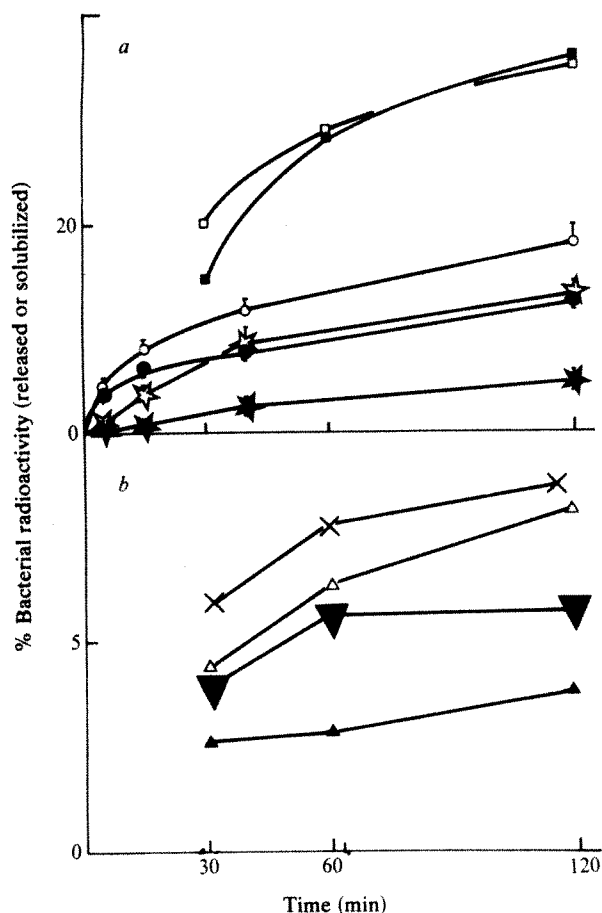


Fig. 3 Solubilization of heat-killed radiolabelled staphylococci by, and release of solubilized radioactivity from, CGD and normal neutrophils. Bacteria were grown in broth containing ^3H -glucose ($25 \mu\text{Ci ml}^{-1}$), ^3H -uridine ($1 \mu\text{Ci ml}^{-1}$) or ^{14}C -phenylalanine ($0.2 \mu\text{Ci ml}^{-1}$) (Radiochemical Centre) and then washed, killed by heating in a boiling water bath for 10 min, rewashed and opsonized with IgG⁴¹. Radiolabelled bacteria were mixed with cells (1:1) and rapidly shaken at 37°C . Solubilization was measured by lysing the cells with distilled water, centrifuging at $8,000g$ for 10 min, homogenizing the pellet in distilled water recentrifuging and counting the radioactivity in the supernatants. Release of solubilized radioactivity was examined by centrifuging the cells and bacteria at $7,000g$ for 1 min after various incubation times, and determining the radioactivity in the supernatant. *a*, Solubilization of uridine (mean of two patients, $\blacksquare$, and controls, $\square$) and glucose (mean $\pm$ s.e. of four patients, $\bullet$, and controls, $\circ$) and release of the solubilized ^3H -glucose label (mean $\pm$ s.e. of four patients, $\star$, and controls, $\star$). *b*, Solubilization of ^{14}C -phenylalanine by cells from two patients ($\blacktriangle$) and controls ($\triangle$). Methylamine (10 mM) enhanced solubilization by both the CGD ($\blacktriangledown$) and control ($\times$) cells.

microscopic examination. After about 2 min the vacuoles in normal, but not in CGD, cells began to swell and became much larger than the enclosed bacterium (Fig. 4). The mean vacuolar volume in normal cells, as calculated from the cross-sectional area of vacuoles on electron micrographs, was approximately $15 \mu\text{m}^3$ after 8 min and $22 \mu\text{m}^3$ after 16 min, in comparison with 4.0 and $4.5 \mu\text{m}^3$ in the patients' cells after the same incubation times (see Fig. 4 legend).

The abnormally small size of phagocytic vacuoles in CGD^{8,31} has been attributed to a reduction in the release of granule contents into the vacuole as a result of defective degranulation. However, as described here and by others³², degranulation is not defective in CGD and although the enlarged vacuoles seen in normal cells appear swollen, they are not obviously distended with electron-dense granule contents. We propose instead that phagocytic vacuoles in normal cells enlarge as a result of the attraction of water by osmotically active digestion products of

the ingested bacteria, and that the vacuoles in CGD cells are much smaller because, as shown in Fig. 3, digestion within their vacuoles is impaired. This hypothesis is supported by our findings that even after incubation times as long as 1 h there is virtually no swelling of vacuoles by normal cells after the phagocytosis of indigestible latex particles opsonized with the same IgG as the bacteria, and that in the presence of methylamine, which enhances bacterial digestion, the mean volume of the patients' vacuoles increased from a calculated 4.5 to $15 \mu\text{m}^3$ ($P < 0.005$, see Fig. 4 legend).

We suggest that the normal sequence of events following phagocytosis of staphylococci is as follows. After entry of the bacteria, the vacuole oxidase system elevates the intravacuolar pH via a cytochrome b_{-245} mediated⁹⁻¹¹ transport of electrons from the cytosolic to the luminal surface. The resultant reduction of molecular oxygen to O_2^{2-} would consume protons during the formation of H_2O_2 (refs 13, 14), raising the intravacuolar pH. The raised pH facilitates killing and bacteriolysis by granule proteins³³⁻³⁶ following which the pH of the vacuoles is reduced to optimize the activities of hydrolases and other proteins with acid pH optima³⁷. In normal vacuoles the generation of low-molecular-weight hydrolysis products is evidently sufficiently rapid to swell osmotically the vacuoles containing the disintegrating organisms. The vacuole enlarges until it comes into contact and fuses with the plasma membrane, releasing the partially digested bacterium and solubilized products, possibly with immunogenic potential, into the extracellular environment. In CGD, however, the absence of the oxidase results in a rapid fall in vacuolar pH inhibiting the killing mechanisms, although not before they have disposed of most of the ingested organisms, and impairing subsequent microbial digestion and its effects on vacuolar morphology. Impaired digestion could account for the lipochromic material that accumulates within histiocytes in the tissues of these patients^{38,39}.

Differences in the control of the pH of endocytic vacuoles might be one of the factors responsible for variations in the susceptibility of individuals to infection and could be manipulated therapeutically⁴⁰.

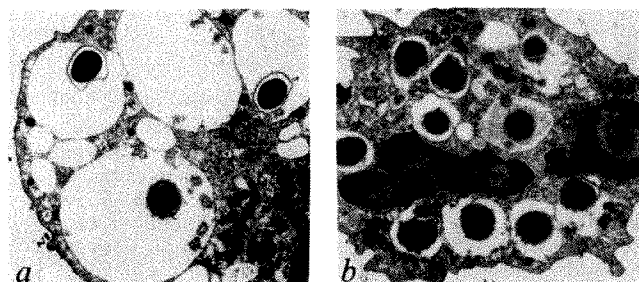


Fig. 4 Transmission electron micrographs of neutrophils from *a*, a normal subject, and *b*, a male patient with CGD ($\times 5,600$). Samples were fixed and processed 16 min after exposure to opsonized staphylococci. Phagocytic vacuoles in normal cells were found to be enlarged and to contain bacterial debris whereas those in CGD cells were much smaller and the bacteria appeared undegraded. Prints were made of electron micrographs of 10 randomly selected cells at each time interval and between 20 and 40 phagocytic vacuoles containing bacteria that appeared to have been sectioned through their region of maximum diameter were excised and weighed. The samples were coded and only identified after the conclusion of the measurements. Cross-sectional areas were determined by comparing the weights of excised vacuoles with those of intact bacteria (radius $\sim 0.6 \mu\text{m}$)⁴³. The cross-sectional areas of the patients' vacuoles were significantly smaller than those of the control (2.8 ± 0.2 and 2.9 ± 0.2 compared with 6.5 ± 1.2 and 8.3 ± 1.0 times that of the bacteria at 8 and 16 min respectively, $P < 0.005$ and $P < 0.001$)²⁵. The cells of another patient were examined and although not quantified, the vacuoles were also obviously much smaller than those in the paired control cells.

We thank the Wellcome Trust and MRC for financial support, Drs David Fellingham and Richard Whitton for helpful advice, and Professors D. Weatherall and M. Greaves for allowing us to study patients under their care.

Received 4 September 1980; accepted 10 February 1981.

1. Baldrige, C. W. & Gerard, R. W. *Am. J. Physiol.* **103**, 235–236 (1933).
2. Sbarra, A. J. & Karnovsky, M. L. *J. biol. Chem.* **234**, 1355–1362 (1959).
3. Selvaraj, R. J. & Sbarra, A. J. *Nature* **211**, 1272–1276 (1966).
4. Mandell, G. L. *Infect. Immun.* **9**, 337–341 (1974).
5. Holmes, B., Page, A. R. & Good, R. A. *J. clin. Invest.* **46**, 1422–1432 (1967).
6. Quie, P. G., White, J. G., Holmes, B. & Good, R. A. *J. clin. Invest.* **46**, 668–679 (1967).
7. Segal, A. W. & Peters, T. J. *Q. J. Med.* **47**, 213–220 (1978).
8. Klebanoff, S. J. & Clark, R. A. in *The Neutrophil*, 641–709 (North-Holland, Amsterdam, 1978).
9. Segal, A. W. & Jones, O. T. G. *FEBS Lett.* **110**, 111–114 (1980).
10. Segal, A. W. & Jones, O. T. G. *Nature* **276**, 515–517 (1978).
11. Segal, A. W. & Jones, O. T. G. *Biochem. J.* **180**, 33–44 (1979).
12. Babior, B. M., Kipnes, R. S. & Curnutte, J. T. *J. clin. Invest.* **52**, 741–744 (1973).
13. Iyer, G. Y. N., Islam, M. F. & Quastel, J. H. *Nature* **192**, 535–541 (1961).
14. Root, R. K., Metcalf, J., Oshino, N. & Chance, B. *J. clin. Invest.* **55**, 945–955 (1975).
15. Tauber, A. I. & Babior, B. M. *J. clin. Invest.* **60**, 374–379 (1977).
16. Green, M. R., Hill, H. A. O., Okolow-Zubrowska, M. J. & Segal, A. W. *FEBS Lett.* **100**, 23–26 (1979).
17. Babior, B. M., Curnutte, J. T. & Kipnes, R. S. *J. Lab. clin. Med.* **85**, 235–244 (1975).
18. Klebanoff, S. J. *Semin. Hemat.* **12**, 117–142 (1975).

19. Tan, J. S., Watanakunakorn, C. & Phair, J. P. *J. Lab. clin. Med.* **78**, 316–322 (1971).
20. Berlin, R. D., Fera, J. P. & Pfeiffer, J. R. *J. clin. Invest.* **63**, 1137–1144 (1979).
21. Segal, A. W. & Allison, A. C. *Ciba Fdn Symp.* **65**, 205–223 (1979).
22. Jacques, Y. V. & Bainton, D. F. *Lab. Invest.* **39**, 179–185 (1978).
23. Ohkuma, S. & Poole, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3327–3331 (1978).
24. Mandell, G. L. *Proc. Soc. exp. Biol. Med.* **134**, 447–449 (1970).
25. Armitage, P. in *Statistical Methods in Medical Research*, 281–282, 289–291 (Blackwell, Oxford, 1971).
26. Segal, A. W., Dorling, J. & Coade, S. *J. Cell Biol.* **85**, 42–59 (1980).
27. Barrett, A. J. & Heath, M. F. in *The Lysosomes. A Laboratory Handbook* (ed. Dingle, T. T.) 19–145 (North-Holland, Amsterdam, 1977).
28. Gorin, G., Wang, S. F. & Papapavlou, L. *Ann. Biochem.* **99**, 113–127 (1971).
29. Bretz, U. & Baggiolini, J. *Cell Biol.* **63**, 251–269 (1974).
30. Davies, P., Page, R. C. & Allison, A. C. *J. exp. Med.* **139**, 1262–1282 (1974).
31. Holmes, B., Quie, P. G., Windhorst, D. B. & Good, R. A. *Lancet* **i**, 1225–1228 (1966).
32. Stossel, T. P., Root, R. K. & Vaughan, M. *New Engl. J. Med.* **286**, 120–123 (1972).
33. Skarnes, R. C. & Watson, D. W. *J. exp. Med.* **104**, 829–845 (1956).
34. Skarnes, R. C. & Watson, D. W. *Bact. Rev.* **21**, 273–294 (1957).
35. Skarnes, R. C. *Nature* **216**, 806–808 (1967).
36. Odeberg, H. & Olsson, I. *J. clin. Invest.* **56**, 1118–1124 (1975).
37. Hirsch, J. G. *J. exp. Med.* **103**, 613–621 (1956).
38. Landing, B. H. & Shirkey, H. S. *Pædiatrics* **20**, 431–438 (1957).
39. Clawson, C. C., Rodey, G. E. & Good, R. A. *Lab. Invest.* **22**, 294–300 (1970).
40. de Duve, C. et al. *Biochem. Pharmac.* **23**, 2495–2531 (1974).
41. Segal, A. W. & Coade, S. B. *Biochem. biophys. Res. Commun.* **84**, 611–617 (1978).
42. Boyum, A. *Scand. J. clin. Lab. Invest.* **21**, Suppl. 97, 31–50 (1968).
43. Luria, S. E. in *The Bacteria* (eds Gunsalus, I. C. & Stamer, R. Y.) **1**, 8 (Academic, New York, 1960).

Deficiency in cells expressing terminal transferase in autoimmune (motheaten) mice

Kenneth S. Landreth*, Kathleen McCoy†, James Claggett‡, F. J. Bollum§ & Cornelius Rosse*

Departments of *Biological Structure, †Microbiology/Immunology, and ‡Periodontics and Microbiology/Immunology, University of Washington School of Medicine, Seattle, Washington 98195, USA
§Department of Biochemistry, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20014, USA

The extensive breakdown of immune homeostasis in the motheaten mouse (*me/me*) has been ascribed to a single gene defect on chromosome 6 (ref. 1). These mice develop skin lesions within the first week of life, do not thrive, and die within the first 3–8 weeks². There is severe hypergamma-globulinaemia³ with multiple species of circulating autoantibody^{3–5} and deposition of immune complexes in the thymus, skin, lungs and kidneys^{1,3}. A single gene defect producing such catastrophic results may provide an important model for understanding autoimmune phenomena. We report here a virtual absence of terminal deoxynucleotidyl transferase-positive (TdT⁺) cells in the bone marrow, thymus and spleen of motheaten mice. TdT is a DNA polymerase which has the unique capacity to polymerize nucleotides in the absence of template direction⁶. Although no *in vivo* biological function of this enzyme has been established, its unique appearance in the bone marrow and thymus of adult mammals and its *in vitro* biochemical activity⁶ have led to a proposed role for TdT in the somatic diversification of lymphocytes⁷. Bone marrow TdT⁺ cells have been shown to belong to both T⁸ and B^{9,10} cell populations and may also include precursor cells common to these lineages¹¹. Although the role of TdT in the acquisition of appropriate T- and B-cell specificities is not known, our results are the first to correlate the virtual absence of TdT⁺ cells with a severe autoimmune syndrome. We investigated the level of TdT⁺ cells in neonatal *me/me* mice and their normal littermates and the susceptibility of TdT⁺ cells to circulating autoantibody in motheaten mouse serum.

Bone marrow, thymus and spleen cells obtained from 21–25-day-old C3HeB/FeJ *me/me* mice, their normal littermates and C3HeB/FeJ wild-type mice were stained for TdT with immunoadsorbent-purified rabbit anti-calf TdT, previously shown to have good cross-reactivity with mouse TdT¹². The

TdT-specific reagent was visualized with fluorescein isothiocyanate (FITC)-conjugated F(ab')₂ goat anti-rabbit IgG. We found an extremely low incidence of TdT⁺ cells in both bone marrow and thymus of *me/me* mice compared with normal littermates and wild-type mice (Table 1). Most of the motheaten mice examined had no TdT⁺ cells, but when present, the numbers of these cells seemed to be inversely correlated with the severity of disease as judged by body weight and skin lesions (unpublished data).

As *me/me* mice possess circulating IgM autoantibodies to thymocytes^{1,5,13}, we tested whether or not motheaten mouse serum had autoantibody to TdT⁺ cells in the lymphoid tissues. Serum was collected from 10 *me/me* mice and 10 normal littermates between 21–25-days old, and incubated with C3HeB/FeJ wild-type bone marrow and thymus cells (Table 2). The binding of *me/me* serum IgM to normal bone marrow and thymus cells was visualized with tetramethylrhodamine isothiocyanate (TRITC)-conjugated specific goat anti-mouse μ (ref. 14) and TdT was stained as already described. As shown in

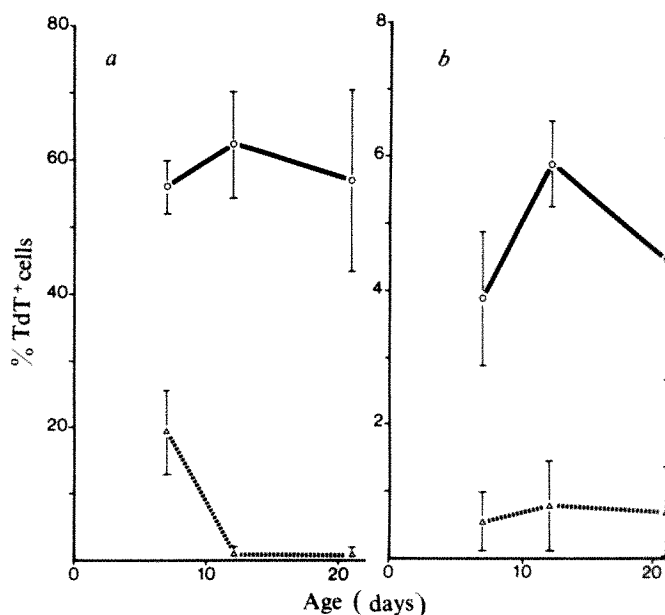


Fig. 1 The incidence of TdT⁺ cells in the thymus (a) and bone marrow (b) of *me/me* mice (Δ) and their normal littermates (○). Cells were collected and stained as described in Table 1. Over 1,000 cells were counted in cytocentrifuge preparations from each of five individual mice; values plotted are mean ± s.d.

† Present address: Sloan-Kettering Institute for Cancer Research, 145 Boston Post Road, Rye, New York 10580, USA.

Table 2, motheaten mice possess circulating IgM autoantibodies that bind to $64.28 \pm 10.37\%$ of normal thymocytes and to $15.38 \pm 7.54\%$ of bone marrow cells; IgM in normal littermate serum samples did not bind to any cells. In the thymus, most of the cells that bound *me/me* autoantibody were also TdT⁺, in contrast to the bone marrow, where none of the cells that bound autoantibody were TdT⁺.

The virtual absence of TdT⁺ cells in the thymus of *me/me* mice in the terminal stages of the autoimmune disease may be due to cytolysis by circulating autoantibody. In the thymus of normal littermates, $56.2 \pm 2.5\%$ of cells were TdT⁺ by 7 days of age and the incidence of these cells remained almost constant up to day 21 (Fig. 1; compare with ref. 6, Fig. 1), but in *me/me* mice only $19.6 \pm 6.8\%$ of thymus cells were TdT⁺ at 7 days. By day 12 the incidence of TdT⁺ cells in the thymus of *me/me* mice had fallen to $0.45 \pm 0.19\%$ and remained at this barely detectable level up to day 21. In normal littermates $3.89 \pm 1.01\%$ of bone marrow cells were TdT⁺ at day 7, increased to $5.91 \pm 0.67\%$ by day 12 and remained at $4.42 \pm 2.10\%$ on day 21. In *me/me* mouse bone marrow only 0.54 ± 0.45 – $0.76 \pm 0.75\%$ of nucleated cells were TdT⁺ at all time points of observation.

One explanation for the virtual absence of TdT⁺ cells in the thymus of motheaten mice by the age of 3 weeks could be autoimmunolysis of TdT⁺ thymocytes—such cell death would be consistent with the decrease in total thymic cellularity during the first 3 weeks¹. Our data suggest that the primary targets for the autoantibody are the TdT⁺ thymocytes found in the thymus cortex¹². Unlike the thymus, the number of TdT⁺ cells in motheaten marrow was low throughout the observation period. Because no binding of autoantibody to TdT⁺ cells could be demonstrated in the bone marrow of 21–25-day-old wild-type mice, it is difficult to attribute the absence of TdT⁺ cells in the marrow to autoimmunolysis. There are four possible explanations for this: (1) cells in defective bone marrow produce a variant of TdT not recognized by our antibody; (2) TdT⁺ cells

Table 1 Incidence of TdT-positive cells in the thymus, bone marrow and spleen of motheaten and normal mice

	% TdT ⁺ cells		
	Thymus	Bone marrow	Spleen
Wild type	52.50 ± 12.34	3.66 ± 1.49	0.12 ± 0.27
Normal littermates	54.25 ± 19.99	4.21 ± 2.08	4.66 ± 2.13
Motheaten	1.43 ± 1.89	0.72 ± 0.81	0.61 ± 0.54

Motheaten mice on a C3HeB/FeJ background, normal littermates and wild-type C3HeB/FeJ mice (obtained from our own breeding colony or from Jackson Laboratories) were killed by cervical dislocation, the femora removed and the contents flushed at 4°C with Hanks' balanced salt solution (HBSS, Microbiological Associates) which contained 5% heat-inactivated fetal calf serum (FCS, Microbiological Associates) and 0.08% NaN₃. Spleens were placed in HBSS at room temperature and teased to extrude cells. Thymic lobes were minced in cold HBSS. Single cell suspensions of each tissue (>90% viable as assessed by trypan blue exclusion) were washed twice in HBSS. An equal volume of 2% paraformaldehyde in phosphate-buffered saline (PBS, pH 7.2) was added to 1.5 – 2.0×10^6 bone marrow, spleen and thymus cells in 0.25 ml HBSS. The cell suspension was kept on ice for 15 min followed by cytocentrifugation (Shandon Scientific) on to clean glass slides at 500 r.p.m. for 3 min. Slide preparations were air dried and immediately fixed in absolute methanol for 20 min at 4°C, then rehydrated through three changes of PBS containing 0.05% NaN₃ at 4°C and dried, excluding the area of the fixed cells which remained wet throughout the staining procedure. Monospecific rabbit anti-TdT (10 µl, Lot A02; 30 µg ml⁻¹ antibody, 100 µg ml⁻¹ BSA) was applied and incubated for 30 min at room temperature in a humidified chamber. Slides were washed three times in excess PBS and incubated a second time with 10 µl of FITC-F(ab')₂ anti-rabbit IgG (10 µg ml⁻¹ in PBS with 100 µg ml⁻¹ of BSA). Slides were finally washed in excess PBS (0.05% NaN₃) for 6–10 h, mounted in Tris-buffered glycerol, pH 9.5, and counted on a Leitz Ortholux epifluorescent microscope. Over 1,000 consecutive cells were counted from each of five individual mice; values are mean $\pm$ s.d.

Table 2 Percentage of bone marrow and thymus cells of wild-type C3HeB/FeJ mice that bind autoantibodies

	Normal littermate serum ⁺	<i>me/me</i> serum ⁺	<i>me/me</i> serum ⁺ and TdT ⁺
Bone marrow	0	15.38 ± 7.54	0
Thymus	0	64.28 ± 10.37	52.50 ± 17.3

To assay circulating autoantibodies in *me/me* mice, bone marrow and thymus cells from wild-type C3HeB/FeJ mice were first incubated with intact or F(ab')₂ goat anti-mouse μ specific antibody for 30 min on ice to block μ determinants on the cell surface. Cells were then washed three times in HBSS (5% FCS, 0.08% NaN₃). Blocking efficiency was checked by incubation of these cell samples with TRITC-conjugated F(ab')₂ goat anti-mouse μ ¹⁴ for 30 min on ice followed by two washes of the cells with HBSS. This procedure routinely reduced detection of μ ⁺ cells to <0.5% in the marrow. One aliquot of the cell samples was mixed with normal mouse serum (diluted 1:5 in HBSS) and another aliquot with *me/me* mouse serum (diluted 1:5 in HBSS) and both were incubated for 30 min on ice. Both cell aliquots were then washed twice in HBSS followed by incubation with TRITC-F(ab')₂ goat anti-mouse μ as described. Washed cell aliquots were fixed in paraformaldehyde as described in Table 1 legend, cytocentrifuged on to clean glass slides and stained for TdT (see Table 1). Cell preparations were mounted in Tris-buffered glycerol, pH 7.0, and counted on a Leitz Ortholux epifluorescent microscope with appropriate filtration for scoring of both FITC-labelled TdT and TRITC-labelled IgM autoantibodies on the same slide. Cell counts represent >1,000 cells counted for each of 10 preparations; mean $\pm$ s.d.

are normally thymus derived; (3) TdT⁺ cells are rapidly exported from motheaten marrow to the involuting thymus or (4) there is a defective production of TdT⁺ bone marrow cells in these mutant mice. The first of these explanations seems unlikely because it requires postulating a species of TdT in *me/me* bone marrow that is antigenically distinct from TdT in *me/me* thymus or from that in normal mouse bone marrow and thymus cells. The post-thymic origin of TdT⁺ bone marrow cells is also considered unlikely because such cells are present in normal numbers in nude mouse bone marrow¹⁵. Present data do not allow a choice between the rapid export and deficient production of TdT⁺ cells. However, even if the paucity of TdT⁺ cells in the bone marrow is due to accelerated migration to the thymus, this explanation would not account for the absence of those TdT⁺ cells in the marrow which are committed to the B-cell lineage, because such cells are unlikely to migrate to the thymus.

The gross abnormality in immunoglobulin production that characterizes the *me/me* syndrome may well be linked to the defect in TdT⁺ B-lineage cells, whereas other manifestations of the *me/me* syndrome may depend on defects in TdT⁺ T cells of both bone marrow and thymic origin. Defective production of TdT or of TdT⁺ cells in the *me/me* mouse may provide important clues about autoimmunity and the normal biological function of TdT. If the enzyme does play a part in the somatic diversification of the molecules by which lymphocytes recognize antigen, then a failure in its production in the developing cell lineage might result in severely decreased heterogeneity of those molecules and in the antigen recognition capacity of the *me/me* mouse. Twenty-one-day-old *me/me* mice fail to respond to foreign antigenic stimulation by either specific antibody production^{3,13,16} or the generation of cytotoxic killer cells¹⁶, yet they exhibit hypergammaglobulinaemia^{1,3,16} and normal numbers of peripheral T cells^{1,5,16}. Preliminary data from our laboratory show that serum immunoglobulins from *me/me* mice have a lower heterogeneity than immunoglobulin from normal littermate serum analysed on two-dimensional polyacrylamide gels after immunoprecipitation of radiolabelled immunoglobulins (D. Putnam, K. McCoy, A. Jones and J. Claggett, unpublished data). These results strengthen the hypothesis that the underlying cause of motheaten pathology may, in part, result from a defect in the production of TdT.

It is intriguing to speculate that a failure in production of TdT⁺ lymphoid precursor cells may be an important defect in autoimmune disease, but other animal models of immune dysfunction need to be examined for changes in TdT-bearing cell populations during the progression of disease. Immunodeficient nude mice have near normal bone marrow TdT⁺ populations that develop early in neonatal life¹⁷ in the virtual absence of TdT⁺ thymocytes. The *nu/nu* mouse defect is primarily related to T cells whereas the *me/me* mutation affects both T and B functions.

This work was supported by USPHS grants CA-20823, CA-23262 and NIH grant DE-04887.

Received 9 December 1980; accepted 19 January 1981.

- Green, M. C. & Shultz, L. D. *J. Hered.* **66**, 250 (1975).
- Kincade, P. W. in *Immune Defects of Laboratory Animals* (eds Gershwin, M. E. & Merchant, B.) (Plenum, New York, in the press).
- Shultz, L. D. & Green, M. C. *J. Immun.* **116**, 936 (1976).
- Shultz, L. D. & Zurier, R. B. in *Genetic Control of Autoimmune Disease* (eds Rose, N. R., Bigazzi, P. E. & Warner, N. L.) 229 (Elsevier, New York, 1978).
- Davidson, W. R., Morse, H. C., Sharrow, S. O. & Chused, T. M. *J. Immun.* **122**, 884 (1979).
- Bollum, F. J. *Blood* **54**, 1203 (1979).
- Baltimore, D. *Nature* **248**, 409 (1974).
- Silverstone, A., Cantor, H., Goldstein, G. & Baltimore, D. *J. exp. Med.* **144**, 543 (1976).
- Janossy, G. *et al. J. Immun.* **123**, 1525 (1979).
- Landreth, K. S., Rosse, C. & Clagett, J. *Fedn Proc.* **39**, 806 (1980).
- Goldschneider, I., Metcalf, D., Mandel, T. & Bollum, F. J. *J. exp. Med.* **152**, 438 (1980).
- Goldschneider, I., Gregoire, K. E., Barton, R. W. & Bollum, F. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 734 (1977).
- Shultz, L. D., Sidman, C. L. & Unanue, E. R. in *Animal Models of Comparative and Developmental Aspects of Immunity and Disease* (eds Gershwin, M. E. & Cooper, E. L.) 260 (Pergamon, New York, 1979).
- Landreth, K. S., Rosse, C. & Clagett, J. (in preparation).
- Hutton, J. J. & Bollum, F. J. *Nucleic Acids Res.* **4**, 457 (1977).
- Sidman, C. L., Shultz, L. D. & Unanue, E. R. *J. Immun.* **121**, 2392, 2399 (1978).
- Sasaki, R., Bollum, F. J. & Goldschneider, I. *J. Immun.* (in the press).

Similarity of teleocidin B and phorbol ester tumour promoters in effects on membrane receptors

Kazuo Umezawa*, I. Bernard Weinstein†, Ann Horowitz‡, Hirota Fujiki‡, Taijiro Matsushima* & Takashi Sugimura‡

* Institute of Medical Sciences, University of Tokyo, Tokyo-108, Japan

† Cancer Center/Institute of Cancer Research, Columbia University College of Physicians and Surgeons, New York 10032, USA

‡ National Cancer Center Research Institute, Tokyo 104, Japan

The tumour promoter 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA) and structurally related phorbol esters effect changes in avian and mammalian cell cultures that mimic transformation by oncogenic viruses or chemical carcinogens and the inhibition or induction of various types of differentiation (for review see refs 1–3). Unlike initiating carcinogens, which seem to act by binding covalently to cellular DNA, the primary site of action of the phorbol ester tumour promoters seems to be the cell membrane^{2,3}; indeed, specific high-affinity saturable receptors for phorbol esters have been identified in cell membranes^{4,5}. The recently discovered class of tumour promoters, the teleocidins⁶, are as potent as TPA in the induction of ornithine decarboxylase in mouse skin, the inhibition of differentiation of Friend erythroleukaemia cells, the induction of HL-60 cell adhesion⁶ and the promotion of tumours on mouse skin⁷. As the teleocidins are structurally unrelated to the phorbol esters, we set out to determine their effects on cell membranes and receptors. We found that in rodent cell cultures, teleocidin B and dihydroteleocidin B have effects similar to those of TPA and that, at nanomolar concentrations, teleocidin inhibits the binding of phorbol esters to cell-surface receptors, which suggests that the action of both classes of compounds may be mediated by the same or a similar receptor system.

TPA enhances the transformation of the mouse embryo cell line C3H 10T1/2 cells on first exposure to radiation or chemical carcinogens^{8–10}. It was of interest, therefore, to see if TPA and the teleocidins exert similar biochemical effects in these cells.

An early membrane response to TPA in mouse skin and certain cell cultures is induction of arachidonic acid release from membrane phospholipids, associated with the synthesis of the arachidonic acid products prostaglandin E₁ and F_{2α} (refs 11–13). In Fig. 1, C3H 10T1/2 cells were prelabelled with ³H-arachidonic acid in conditions in which >95% of the acid is incorporated in phospholipids¹². The subsequent addition of TPA, teleocidin B or dihydroteleocidin B (100 ng ml⁻¹) to these cultures led to the rapid release of ³H-arachidonic acid into the media (Fig. 1a), reflecting increased phospholipid degradation. As with TPA, the teleocidin-induced release of arachidonic acid was also associated with increased synthesis of prostaglandins E₁ and F_{2α} (data not shown here). Concentration curve studies indicated that the optimum concentrations of teleocidin B and dihydroteleocidin B were ~10 ng ml⁻¹ and that these agents were actually more potent than TPA in inducing arachidonic acid release (Fig. 1b). The exposure of avian and mammalian cells to TPA causes a rapid enhancement in the uptake of 2-deoxyglucose^{14,15} and certain other nutrients^{2,3}; we found that teleocidin B and dihydroteleocidin B stimulated 2-deoxyglucose uptake by C3H 10T1/2 cells to levels comparable with those obtained for TPA when all three agents were tested at 100 ng ml⁻¹ concentrations (data not shown).

An additional early and specific response of cells to TPA is inhibition of binding of epidermal growth factor (EGF) to its cell-surface receptors^{16–19}. Although this does not seem to be due to direct binding of TPA to the receptor^{17–19}, it does provide a highly sensitive marker of membrane changes induced by the

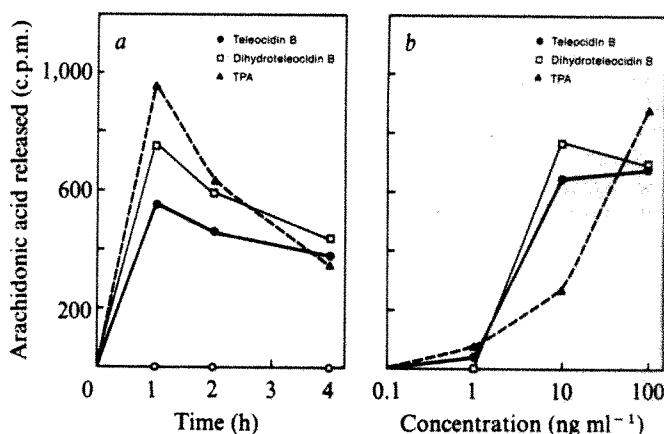


Fig. 1 Arachidonic acid and prostaglandin release were assayed as described elsewhere. Phospholipids of C3H 10T1/2 cells were prelabelled by incubation of 4×10^5 cells per plate with 1.25 μ Ci per ml of ³H-arachidonic acid (62.2 Ci mmol⁻¹, NEN) for 24 h. The prelabelled cells were then washed and incubated in 2 ml of serum-free Dulbecco's minimal essential medium (DMEM) containing either solvent or the indicated chemicals at 37 °C. At the indicated times the media were removed, acidified and extracted with ether. After evaporation of the ether, the sample was solubilized in 0.2 ml of chloroform. A portion (0.04 ml) was applied to silica gel TLC plates with nonradioactive markers, developed in 2,6-dimethylheptanone/glacial acetic acid/0.9% NaCl (80:40:6), and the plate scraped and assayed for radioactivity. Values are means of duplicate samples which agreed within 10%. Two additional experiments gave similar results. *a*, Samples were incubated with 30 ng per ml of TPA (▲), teleocidin B (●), dihydroteleocidin B (□), or the dimethyl sulphoxide (DMSO) solvent (○) for the times indicated. *b*, Samples were incubated with the indicated concentrations of TPA (▲), teleocidin B (●), dihydroteleocidin B (□), or DMSO (○) for 1 h at 37 °C. With the DMSO solvent control there was negligible release of radioactivity, presumably because spontaneous phospholipid turnover is reduced in the serum-free incubation system (ref. 12 and unpublished results).

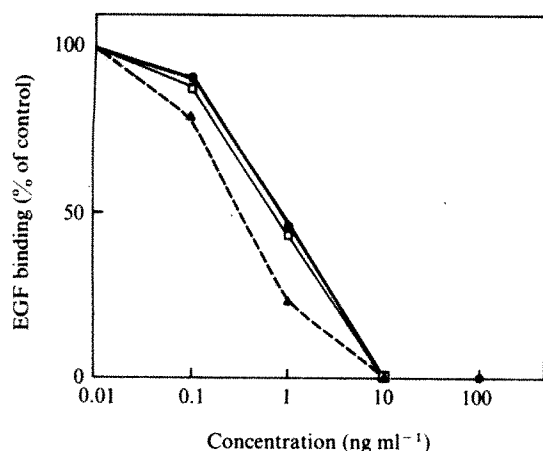


Fig. 2 Inhibition of EGF binding: 2×10^6 C3H 10T1/2 cells were incubated with 0.05 μ Ci (0.2 ng) of ^{125} I-EGF (120 mCi per mg, Collaborative Research) in 2 ml of serum-free DMEM with the indicated concentration of TPA ($\blacktriangle$), teleocidin B ($\bullet$), or dihydroteleocidin B ($\square$) for 2 h, washed and assayed for cell-bound radioactivity as described elsewhere¹⁷. The control value was 940 c.p.m. per dish and is expressed as 100%. Values are means of duplicate samples which agreed within 10%. Two additional experiments gave similar results.

phorbol ester tumour promoters. Figure 2 indicates that the addition of 0.1–10 ng ml^{-1} of TPA, teleocidin B or dihydroteleocidin B simultaneously with the addition of ^{125} I-EGF led to a marked inhibition of ^{125} I-EGF binding to C3H 10T1/2 cells. All three compounds also inhibited ^{125} I-EGF binding in a line of rat embryo cells (FRE-8D) and in HeLa cells (data not shown). In C3H 10T1/2 and FRE-8D cells, TPA was rather more potent than the teleocidins in inhibiting ^{125} I-EGF binding whereas the three compounds had similar potencies for HeLa cultures.

Several aspects of the action of the phorbol esters have led to the suggestion that they act by usurping cell membrane receptors normally occupied by an endogenous growth factor^{1,16}. Using ^3H -phorbol-12,13-dibutyrate (PDBu) to avoid non-specific binding, Blumberg⁴ has in fact recently demonstrated that membrane fractions from avian and mouse cells contain saturable high-affinity receptors for this compound. A modification of this approach has also demonstrated receptors for ^3H -PDBu on intact monolayers of rat embryo fibroblasts and several other cell types⁵. Figure 3 shows that at 37 °C, concentrations of teleocidin B in the range 0.5–10 ng ml^{-1} were able to produce marked inhibition of the specific binding of ^3H -PDBu to rat embryo fibroblast monolayer cultures; 50% inhibition was achieved with $\sim 2.5 \text{ ng ml}^{-1}$, which is equivalent to about 10^{-9} M . Teleocidin was more potent than TPA in inhibiting ^3H -PDBu binding, because $\sim 3.8 \text{ ng ml}^{-1}$ of TPA were required to produce 50% inhibition (Fig. 3). The ID_{50} value for PDBu was 23 ng ml^{-1} (data not shown). No inhibition was observed with as much as 1,000 ng ml^{-1} of phorbol, which lacks tumour promoting activity, nor was inhibition obtained with EGF, presumably because it occupies separate receptors^{4,5,17}. In additional studies (not reported here) of ^3H -PDBu binding, carried out at 4 °C rather than 37 °C, teleocidin B was also a more potent inhibitor than TPA. Thus, in contrast to the TPA inhibition of EGF binding^{17–19}, teleocidin B inhibition is not highly temperature dependent. These results provide further evidence that teleocidin B competes directly with ^3H -PDBu for receptor binding.

There is evidence that the TPA induction of arachidonic acid release, TPA enhancement of 2-deoxyglucose uptake and TPA inhibition of EGF-receptor binding are direct membrane effects that do not require *de novo* RNA synthesis^{2,11,12,17}. Because of their similar time courses we would assume that this is also true

for the effects of teleocidin on these membrane-related properties. These results, together with our evidence that teleocidin B is a potent and apparently direct inhibitor of the binding of PDBu to its cell-surface receptors, suggest that the teleocidins might act as tumour promoters and induce changes in cell culture by binding to the same membrane-associated receptors as those used by TPA.

We emphasize that although teleocidin B was about 10 times more potent than TPA in inducing arachidonic acid release (Fig. 1) and slightly more potent in inhibiting ^3H -PDBu binding (Fig. 3), it was slightly less potent than TPA in inhibiting EGF-receptor binding (Fig. 2). Because these parameters are quite different and were not all assayed with the same cell cultures, the significance of these quantitative differences is unknown. It is possible that they reflect the existence of subtypes of ^3H -PDBu receptors⁵ that control different cellular effects, but this requires further study.

Teleocidin has an entirely different chemical structure from that of TPA or any of the TPA-related plant diterpenes, but both compounds are somewhat amphipathic. Teleocidin has a peptide-like moiety joined to a lipid portion containing two isoprenes⁶, while TPA contains a somewhat hydrophilic-substituted ring system and a long-chain lipophilic fatty acid residue³. These aspects may provide clues to the molecular details by which the two different classes of compound apparently bind to the same or a similar set of membrane receptors. In addition, the fact that two structurally diverse compounds with potent tumour promoting activity seem to bind to the same cell-surface receptors suggests that these receptors play a part in mediating the biological effects of the compounds.

We thank Drs R. A. Mufson and P. B. Fisher for their advice and Dr M. Takashima for providing dihydroteleocidin B. This research was supported by NCI grant CA-26056, DHEW, an

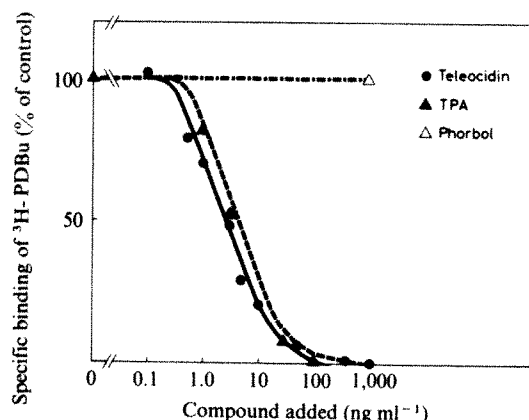


Fig. 3 Inhibition of specific binding of ^3H -PDBu. Rat embryo fibroblast cells (FRE-8D) were grown in 6-cm dishes containing DMEM plus 10% calf serum. The assay was performed on 2×10^6 cells per dish. The cell monolayer was washed once with phosphate-buffered saline (PBS) and then 2 ml of assay buffer (2 volumes DMEM/1 volume PBS, plus 1 mg ml^{-1} bovine serum albumin) were added. The indicated concentrations of teleocidin B ($\bullet$), TPA ($\blacktriangle$) or phorbol ($\triangle$) were added, followed by ^3H -PDBu (6.4 Ci mmol^{-1} , Lifesystems) to a final concentration of 3 nM. The cells were incubated for 30 min at 37 °C to achieve plateau binding, washed 3 times with assay buffer at 0 °C, solubilized in 1 ml of 0.8% Triton X-100, 0.02% EDTA and 0.25% trypsin in PBS, and assayed for radioactivity. Nonspecific ^3H -PDBu binding was assumed to be that obtained in the presence of 50 μM unlabelled PDBu. This value was $\sim 20\%$ of the total binding obtained in the absence of excess PDBu and was subtracted from all experimental data to obtain the values for specific saturable binding. Specific binding of ^3H -PDBu in the control assays was ~ 750 c.p.m. per 10^6 cells (equivalent to $\sim 1.1 \times 10^5$ receptors per cell), with a s.e.m. of $<10\%$. This control value is expressed as 100%. Similar results were obtained in two additional experiments. Further details of this assay will be published elsewhere⁵.

award from the DuPont Company and in part by grants-in-aid for cancer research from the Ministry of Education, Science and Culture and the Ministry of Health and Welfare of Japan.

Received 23 July 1980; accepted 26 January 1981.

- Weinstein, I. B., Wigler, M. & Pietropaolo, C. in *Origins of Human Cancer* (eds Hiatt, H. H., Watson, J. D. & Winston, J. A.) 751–772 (Cold Spring Harbor Laboratory, New York, 1977).
- Weinstein, I. B., Lee, L. S., Fisher, P. B., Mufson, R. A. & Yamasaki, H. *supramolec. Struct.* **12**, 195–208, 1979.
- Slaga, T. J., Sivak, A. & Boutwell, R. K. (eds) *Carcinogenesis Vol. 2* (Raven, New York, 1978).
- Delclos, K. B., Nagle, D. S. & Blumberg, P. D. *Cell* **19**, 1025–1032 (1980).
- Horowitz, A., Greenebaum, E. & Weinstein, I. B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Fujiki, H., Mori, M., Nakayasu, M., Terada, M. & Sugimura, T. *Biochem. biophys. Res. Commun.* **90**, 976–983 (1979).
- Fujiki, H. *et al.* (in preparation).
- Mondal, S., Brankow, D. W. & Heidelberger, C. *Cancer Res.* **36**, 2254–2260 (1976).
- Mondal, S. & Heidelberger, C. *Nature* **260**, 710–711 (1976).
- Kennedy, A., Mondal, S., Heidelberger, C. & Little, J. B. *Cancer Res.* **38**, 439–443 (1978).
- Levine, L. & Hassid, A. *Biochem. biophys. Res. Commun.* **79**, 477–483 (1977).
- Mufson, R. A., DeFeo, D. & Weinstein, I. B. *Molec. Pharmac.* **16**, 569–578 (1979).
- Yamasaki, H., Mufson, R. A. & Weinstein, I. B. *Biochem. biophys. Res. Commun.* **89**, 1018–1025 (1979).
- Driedger, P. E. & Blumberg, P. M. *Cancer Res.* **37**, 3257–3265 (1977).
- Lee, L. S. & Weinstein, I. B. *J. cell. Physiol.* **99**, 451–460 (1979).
- Lee, L. S. & Weinstein, I. B. *Science* **202**, 313–315 (1978).
- Lee, L. S. & Weinstein, I. B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5168–5172 (1979).
- Shoyab, M., DeLarco, J. E. & Todaro, G. J. *Nature* **279**, 387–391 (1979).
- Brown, K. D., Dicker, P. & Rozenfurt, E. *Biochem. biophys. Res. Commun.* **86**, 1037–1043 (1979).

Synaptic facilitation requires paired activation of convergent pathways in the neocortex

A. Baranyi & O. Fehér

Department of Comparative Physiology, Attila József University of Sciences, Közpétfasor 52, Szeged 6726, Hungary

In associative learning, the activated neurones undergo a variety of concomitant functional alterations—increases or decreases of firing activity^{1–4} and modifications of membrane potential or resistance^{5–8} and of synaptic responsiveness^{9–11}. Synaptic transmission which can be strengthened only when there is paired activity in two pathways is of particular interest in relation to mechanisms for associative learning^{12,13}. For the neocortex, there are few observations of the plastic changes, induced by conditioning procedures, in the effectiveness of individual synapses^{14–17}. We now report that various regimes with joint stimulations of convergent excitatory pathways on to intracellularly recorded neurones in the motor cortex of the cat result in synaptic facilitation lasting for up to 30 min.

The experiments were performed on the exposed motor cortex of cats anaesthetized with 35 mg per kg Nembutal using conventional intracellular recording techniques. The surgical procedure, methods of stimulation and recording are reported elsewhere in detail^{15,16}. Synaptic responses were elicited by stimulation of the pyramidal tract, the ventrolateral thalamic nucleus (VL), the homotopic cortical surface (callosal stimuli) and the contralateral radial nerve (somatosensory activation). The membrane input resistance, R_m , was measured by the bridge-balancing method¹⁸. At least two excitatory synaptic inputs on to the intracellularly recorded cell were stimulated: one (test pathway) elicited only mono- or oligosynaptic excitatory postsynaptic potentials (test e.p.s.ps), while the other caused the cell to produce spike potentials (unconditioned stimulus, US) (Fig. 1a). Based on the cellular analogues of classical conditioning^{9,19} the applied conditioning paradigm consisted of (1) habituation, (2) pseudoconditioning, (3) conditioning, and (4) extinction procedures (Fig. 1b). A neurone was regarded as conditionable if the test e.p.s.p. was enhanced in

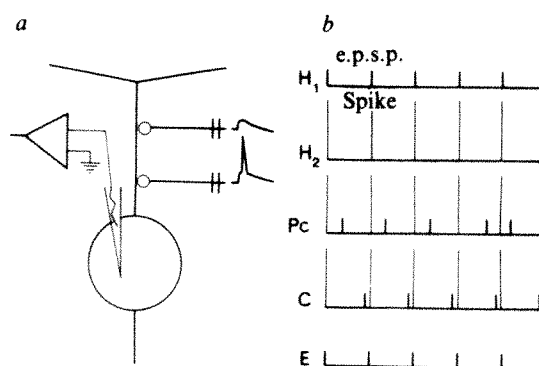


Fig. 1 a, Schematic drawing of the experimental arrangement used for cellular conditioning of neurones in the motor cortex. The stimulation of the test pathway produced an e.p.s.p. while stronger stimuli via another pathway elicited spike potentials in the recorded neurone. b, Illustration of the experimental paradigm. During the habituation procedure 150 test e.p.s.ps (H_1) and then 150 US spikes (H_2) were evoked from the respective inputs at a stimulus frequency of 0.2–1 Hz. Pc, pseudoconditioning procedure when test e.p.s.ps and US spikes were randomized with 0.2–1 Hz frequency within a block of 150 stimuli. C, conditioning procedure with e.p.s.p.–spike stimulus pairs at constant inter-stimulus interval ranging from 0 to 100 ms. E, extinction period when only test stimuli were given at 0.2–1 Hz until the augmented neuronal responses returned to the preconditioning level.

amplitude or in efficacy in evoking spike potentials by the conditioning procedure and if these changes were abolished by the extinction procedure.

In 38 cells (26% of the neurones tested) the efficacy of the test e.p.s.ps was enhanced after 60–90 stimulus pairings and the facilitated synaptic response persisted for periods from 3 to 41 min (mean $\pm$ s.d. 14.5 ± 9.5 min, $n = 38$). Figure 2 shows an example of synaptic facilitation by VL e.p.s.p.–callosal spike repetitive stimulus pairs. In other neurones e.p.s.p.–spike and spike–e.p.s.p. sequences were equally efficient, but only in stimulus pairs with inter-stimulus intervals of <40–60 ms in the former and ~ 100 ms in the latter. Inter-stimulus intervals above 100 ms (periods ranging from 100 to 500 ms were tested) proved ineffective in causing synaptic facilitation. This may be explained by supposing that both US spikes and test e.p.s.ps have an after-effect of about 100 ms in duration and their repetitive overlap is necessary for the facilitatory interaction between them. On the other hand, it could be demonstrated that the observed synaptic facilitation was specific to the pairing procedure, as repeated but unpaired presentation of test e.p.s.ps and US spikes during pseudoconditioning procedures never led to plastic changes (tested in 17 cases, Fig. 2b), and when a previously efficient e.p.s.p.–spike stimulus pair was applied together with a hyperpolarizing current pulse through the recording electrode (when spike generation by US stimuli was prevented), the test e.p.s.p. was not augmented (tested in 13 cases). The amount of hyperpolarization required to prevent action potentials by US stimuli during conditioning was sufficient to abolish the plastic after-effect. Based on the presented data, one may conclude that contingent pairing of the test e.p.s.ps with spike potentials, that is, temporal contiguity of pre- and postsynaptic excitations, is indispensable for the synaptic facilitation, as postulated previously^{20,21}.

Transient facilitations of the test e.p.s.ps lasting for 6–40 s were observed in all 17 neurones with which spike trains of 5–50 Hz frequency and 5–15 s duration were applied as US. This after-effect did not require co-active stimulation of the test and US pathways, because it also occurred when US stimuli were applied on their own. These short-lived effects presumably originated from transient modifications of the ionic micro-environment caused by frequent firing activity^{22–24}. In hippocampal neurones long-term facilitation of both e.p.s.ps and inhibitory postsynaptic potentials has been demonstrated as a

result of repetitive activation with similar parameters^{25,26}. In the motor cortex a considerably higher frequency and longer period of repetitive activation is necessary for long-term facilitation²⁷.

Enhancement of the amplitude of paired test e.p.s.ps was the only consequence of paired conditioning procedures in 21 cells, whereas in the other 17 neurones the synaptic facilitation was accompanied by 3–6 mV membrane depolarization and increases of R_m (Fig. 3). The depolarization presumably did not contribute to the augmentation of the test e.p.s.ps, as a test e.p.s.p. induced during a depolarizing current injection (causing 2–6 mV depolarization) was found in 11 tested cells to be decreased in amplitude or unaffected.

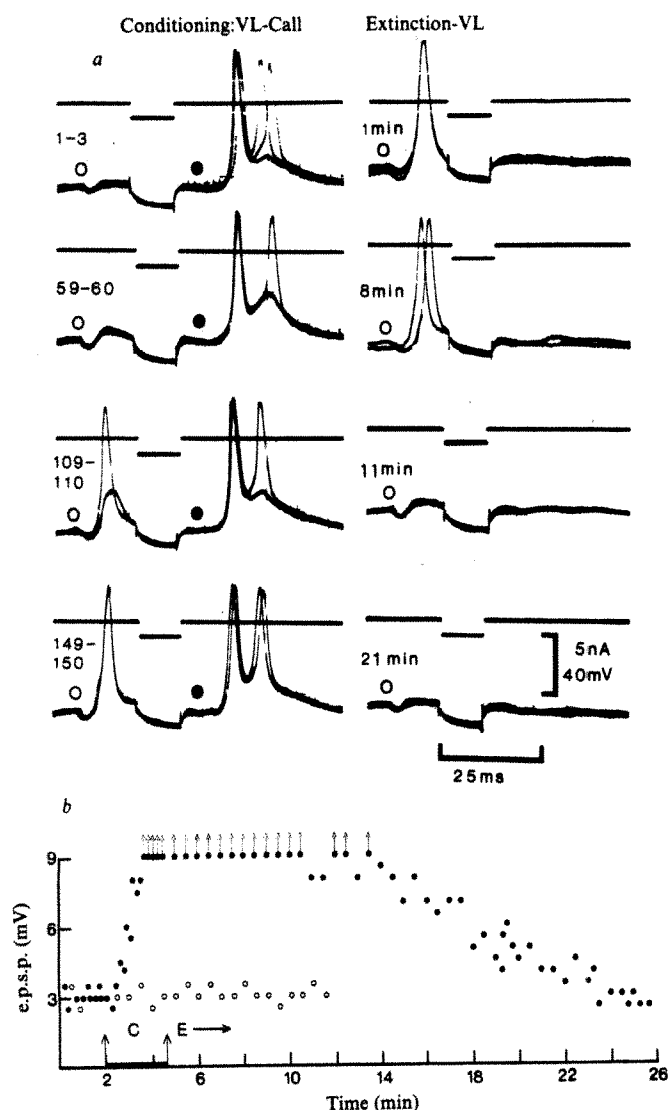


Fig. 2 Facilitation of a test e.p.s.p. by stimulus pairs applied repetitively in a non-pyramidal tract neurone²⁹. Top traces, current passed through the recording microelectrode; bottom traces, intracellular records. The test e.p.s.ps were evoked by stimulation of the VL nucleus (○); US spikes were elicited from the callosal system (Call, ●). The VL e.p.s.p. had a latency of 1.6 ms and was monosynaptic being able to follow 100 Hz frequency of stimulation. *a*, Neuronal responses during the conditioning and extinction procedures repeated once every second with 27 ms inter-stimulus interval. Following the 110th pairing all the test stimuli elicited spike potentials. The enhanced synaptic response returned to the preconditioning level by about the 21st min of the extinction period. R_m was measured every second with 1.2 nA anodal current pulses. No changes in R_m or spontaneous firing were observed during conditioning. *b*, Time course and specificity of facilitation of test e.p.s.p. for pairing procedure. ● (Paired e.p.s.p.) represents an actual data point every 10th second of the conditioning and every half minute of the extinction period. In the unpaired situation (○), only data at each half-minute intervals are presented.

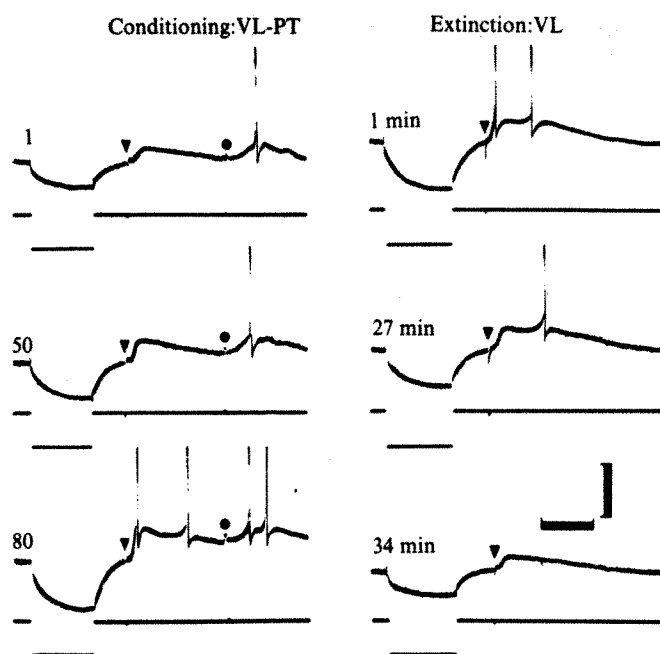


Fig. 3 Conditioned facilitation of a VL test e.p.s.p. (▼) in a non-pyramidal tract neurone with accompanying membrane depolarization and increases in R_m . The US spikes were elicited by pyramidal tract (PT) stimuli (●). Inter-stimulus interval 30 ms, frequency of pairing 0.5 Hz. Top traces, intracellular records; bottom traces, current records. Only the lower part of action potentials are shown because of high amplification. Calibration: 15 mV for intracellular records and 1.66 nA for current records. Time calibration 15 ms. The amplitude of PT spikes was 71 mV. The test e.p.s.ps and PT spikes were monosynaptically evoked (latency, 1.2 and 4.8 ms at 100 Hz stimulus frequency, respectively). During conditioning the preconditioning value of the amplitude of test e.p.s.ps (mean $\pm$ s.d. 4 ± 0.2 mV, $n = 10$) gradually increased and attained maximal facilitation by the 80th pairing, when each VL stimulus evoked two spikes. Note the conditioned decreases of the mean time to peak of VL e.p.s.ps from 2.7 ± 0.1 to 1.7 ± 0.1 ms ($n = 5$). Parallel with the synaptic facilitation, the R_m increased from 5.1 ± 0.3 to 8.8 ± 0.3 M Ω ($n = 10$), while the membrane potential depolarized from 67 ± 2 to 63 ± 2 mV ($n = 10$). All the conditioned changes of test e.p.s.ps and membrane parameters completely disappeared by the 34th min of the extinction period, when only VL stimuli were given alone with 0.5 Hz frequency. None of the above changes occurred during the pseudo-conditioning procedure, when PT and VL stimuli were randomized.

Intracellular observations in the motor cortex of the cat²⁸ and in invertebrate neurones⁵ have suggested that in some conditions an inward calcium current is initiated which leads, via a second messenger system, to membrane depolarization and an elevation in R_m . On the other hand, the R_m increases associated with conditioned increases in firing rate might be due to a decrease in K^+ conductance⁶. Our observations demonstrate that conditioned increases of synaptic efficacy originate from some kind of, as yet unknown, focal postsynaptic processes in the motor cortex without any general modification of the neuronal excitability. Although concomitant alterations of membrane potential or conductance did occur in some of the conditioned neurones, presumably they represent a different type of plastic modification. The mechanism of synaptic facilitation described here requires further examination.

Received 8 April 1980; accepted 28 January 1981.

1. Kandel, E. R. & Spencer, W. A. *Physiol. Rev.* **48**, 65–134 (1968).
2. Berger, T. W. & Thompson, R. F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1572–1576 (1978).
3. Morrel, F. O. In *The Neurosciences, A Study Program* (eds Quarten, G. C., Melnechuk, T. & Schmitt, F. O.) 452–469 (Rockefeller University Press, New York, 1967).
4. Morrel, L. T., Hoepfner, T. J. & Morrel, F. *Science* **204**, 528–530 (1979).
5. Alkon, D. L. *Science* **205**, 810–816 (1979).
6. Woody, C. D. & Black-Cleworth, P. *J. Neurophysiol.* **36**, 1104–1115 (1973).

7. Kandel, E. R. & Tauc, L. *J. Physiol., Lond.* **183**, 287-304 (1966).
8. Hoyle, G. *Neurosci. Res. Prog. Bull.* **17**, No. 4, 577-586 (1979).
9. Kandel, E. R. & Tauc, L. *J. Physiol., Lond.* **181**, 1-47 (1965).
10. Kandel, E. R. & Tauc, L. *Nature* **202**, 145-147 (1964).
11. Voronin, L. L. in *Soviet Research Reports Vol. 2* (ed. Woody, C. D.) (Brain Information Service, University of California, Los Angeles, 1976).
12. McNaughton, B. L., Douglas, R. M. & Goddard, G. V. *Brain Res.* **157**, 277-293 (1978).
13. Gardner-Medwin, A. R. *Trans. Biochem. Soc.* **6**, 841-844 (1978).
14. Voronin, L. L. *Acta neurobiol. exp.* **40**, 1B, 335-370 (1980).
15. Baranyi, A. & Fehér, O. *Expl Brain Res.* **33**, 283-298 (1978).
16. Baranyi, A. & Fehér, O. *Expl Brain Res.* **41**, 124-134 (1981).
17. Tzebelikos, E. & Woody, C. D. *Soc. Neurosci. Abstr.* **4**, 81 (1978).
18. Nelson, P. G. & Frank, K. *J. Neurophysiol.* **30**, 1097-1113 (1967).
19. Kandel, E. R. in *The Neurosciences, A Study Program* (eds Quarten, G. C., Melnechuk, T. & Schmitt, F. O.) 666-689 (Rockefeller University Press, New York, 1967).
20. Hebb, D. O. *The Organization of Behaviour* (Wiley, New York, 1949).
21. Konorski, J. *Conditioned Reflex and Neurone Organization* (Cambridge University Press, London, 1948).
22. Heinemann, U. & Lux, H. D. *Brain Res.* **93**, 63-76 (1975).
23. Koike, H., Mano, N., Okada, Y. & Oshima, T. *Expl Brain Res.* **14**, (1972).
24. Moody, W., Futamachi, K. & Prince, D. A. *Expl Neurol.* **42**, 248-263 (1974).
25. Misgeld, U., Sarvey, J. M. & Klee, M. R. *Expl Brain Res.* **37**, 217-229 (1979).
26. Lynch, G., Gall, C. & Dunwiddie, T. *Prog. Brain Res.* **48**, 113-130 (1978).
27. Bindman, L. J., Lippold, O. C. J. & Milne, A. R. *J. Physiol., Lond.* **286**, 457-477 (1979).
28. Woody, C. D., Swartz, B. E. & Gruen, E. *Brain Res.* **158**, 373-375 (1978).
29. Takahashi, E. *J. Neurophysiol.* **27**, 908-924 (1965).

Response of an insect photoreceptor: a simple log-normal model

Richard Payne & Jonathon Howard

Department of Neurobiology, Research School of Biological Sciences, Australian National University, PO Box 475, Canberra City, ACT 2601, Australia

Photoreceptors respond to light with a graded depolarization or hyperpolarization of the cell membrane. The transduction of light energy to an electrical event, though much studied, remains an outstanding problem in visual research. Previous transduction models¹, involving cascades of chemical reactions, do not fit the responses of insect photoreceptors². It has been proposed that the absorption of light causes the release of an internal transmitter³. We have developed here a simple two-parameter model of transduction based on the assumption that membrane channels are activated when the concentration of internal transmitter reaches a threshold level⁴. The theory predicts that the impulse-response function follows the log-normal curve—this model fits the response of the locust compound eye photoreceptor over the 10-fold range of time scale produced by altering the temperature or the state of light adaptation of the cell.

Photoreceptors in the compound eye of *Locusta migratoria* generate a variable number of superimposed, depolarizing voltage events when stimulated with a brief, weak flash of light⁵. Each event or 'bump' represents an effective quantal absorption of light⁶. The average of a large number of such responses accurately defines the impulse-response function of the receptor. The amplitude of this response increases linearly with increasing flash intensity when the responses have an amplitude of <5 mV and in this linear range, the Fourier transformation of the impulse-response function corresponds to the temporal frequency response of the receptor⁷.

Using glass micropipettes, recordings of electrical potential were made from photoreceptors in an intact eye⁸ and a superfused retina⁹. For the superfused retina we averaged responses from dark-adapted photoreceptors to 1-ms flashes from a green light-emitting diode (Siemens LD57C) at temperatures of the surrounding perfusion fluid between 33 and 17°C. When recording from photoreceptors in an intact eye, a 1-ms test flash from a point light source (also a LED) could be superimposed on steady, adapting illumination from the same source which was

placed on the optical axis of the ommatidium containing the impaled cell.

We attempted to fit our impulse-response functions to the models of Fuortes and Hodgkin¹ and Baylor *et al.*¹⁰. These models are based on the time course of the concentration of the penultimate product of a series of n reactions which proceed sequentially¹ or independently¹⁰. Like Borsellino *et al.*¹¹, we found that the Fuortes and Hodgkin model predicted a response that was too symmetrical about the peak and it also required a range of values of n between 13 and 20 to fit the data from all our cells. The model of Baylor *et al.* fitted the response shape adequately, but required values of n between 20 and 48. The large values for the number of steps required in the reaction chains of these models and the variability from cell to cell made

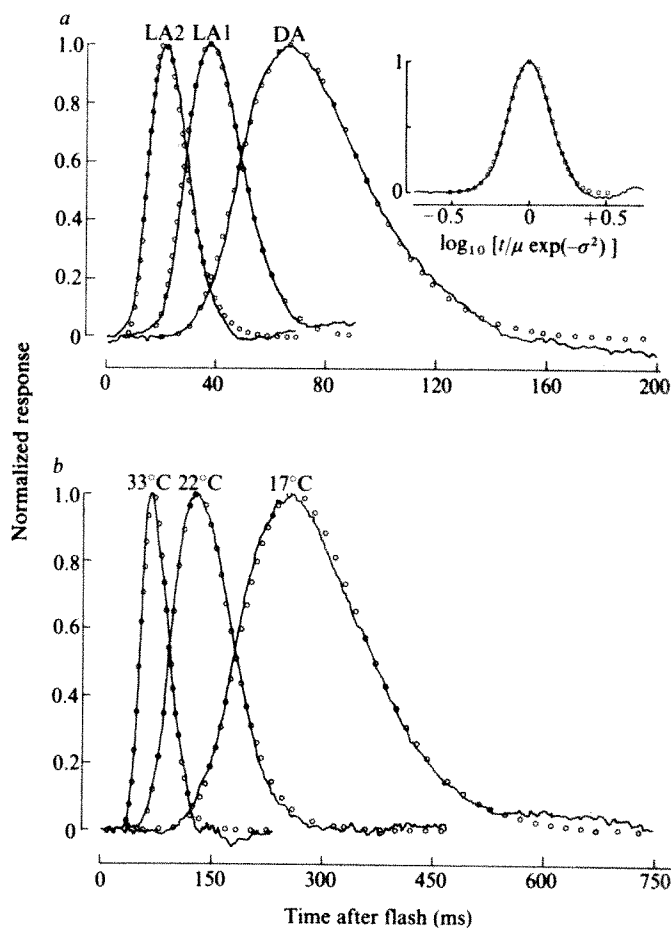


Fig. 1 *a*, Impulse-response functions recorded from a photoreceptor in an intact eye when dark adapted and at two light-adapted states. One hundred presentations of a 1-ms light flash were averaged in each case. The intensities of the flashes and superimposed adapting light were: 14 photons per facet per flash (dark adapted), 42 photons per facet per flash superimposed on an adapting light of 1.3×10^4 photons per facet per s and 4.2×10^5 photons per facet per flash superimposed on an adapting light of 1.3×10^5 photons per facet per s, wavelength 560 nm. Circles show the fits to the data using equation (1) and the following parameters: DA: $\sigma = 0.31$, $\mu = 74$ ms; LA1: $\sigma = 0.27$, $\mu = 41$ ms; LA2: $\sigma = 0.31$, $\mu = 23$ ms. The amplitudes of the responses were 1.5 (LA1) and 2.3 (LA2) mV. Inset, the impulse-response function in the dark-adapted state is replotted on a logarithmic time scale. Equation (1) (circles) predicts that a gaussian curve will have mean $\mu \exp(-\sigma^2)$ and standard deviation σ . *b*, Impulse response functions of a photoreceptor in a perfused eye at three temperatures. Circles show the fit to the data of equation (1) using the following parameter values: 17°C: $\sigma = 0.305$, $\mu = 22$ ms; 22°C: $\sigma = 0.294$, $\mu = 142$ ms; 33°C: $\sigma = 0.247$, $\mu = 71$ ms. Responses shown are averages of 100 presentations of a flash of intensity 1×10^{12} photons $\text{cm}^{-2} \text{s}^{-1}$ and wavelength 560 nm. Amplitudes of the responses, which have been scaled to 1 mV, were 3.1 (17°C), 3.9 (21°C) and 3.8 (33°C) mV.

us doubt their physical applicability. We therefore searched for an alternative model.

We chose to fit our impulse-response functions to the two-parameter log-normal curve (Fig. 1a)

$$V(t) = sI\Delta t \cdot t^{-1} \cdot \exp[-(\log(t/\mu))^2/2\sigma^2] \quad (1)$$

where I is the intensity of the flash of duration Δt delivered at $t = 0$ and s is the sensitivity of the cell. σ is the standard deviation and μ the mean of the transformed gaussian curve.

An adapting background of intensity 3×10^4 photons per facet per s reduced the value of μ for responses from cells in the intact eye from 60 ± 18 ms ($n = 6$) in the dark-adapted state to 24 ± 2 ms when light adapted. The corresponding values of σ , 0.30 ± 0.01 and 0.32 ± 0.02 , showed no significant change (for example, see Fig. 1a). A large decrease in the sensitivity of the response also occurred during light adaptation as noted by others^{12,13}. Lowering the temperature of the superfused preparation from 33 to 21 °C resulted in an increase in μ from 70 ± 22 ms ($n = 7$) to 160 ± 14 ms. A marginal increase in σ from 0.28 ± 0.02 to 0.32 ± 0.02 also occurred (for example, Fig. 1b). Temperature and light adaptation therefore significantly affect only one temporal parameter in our model (μ), corresponding to a change in the time scale of the response. We propose that this parameter represents the rate constant of a single transduction process, being sensitive to both light adaptation and temperature.

The log-normal distribution has a general application to the analysis of biological data with a large variance¹⁴. In particular, Gaddum¹⁵ and Bliss¹⁶ have used the distribution to describe the variability, from subject to subject, in the critical dose of a drug that will just cause a reaction. We propose an analogous physical model of phototransduction. The model assumes that, following the absorption of light, the strength of a chemical signal increases. The broad impulse-response arises from a scatter in the detection of threshold levels of this signal that trigger the activation of depolarizing ionic channels in the photoreceptive membrane.

Specifically, we assume that: (1) a photoisomerization opens several ionic channels in the vicinity of the activated rhodopsin molecule through the action of an internal transmitter. A given channel opens for a brief time when the concentration of transmitter exceeds a threshold value, α . (2) The threshold

value, α , is normally distributed from channel to channel. (3) The time course of local transmitter concentration following a photoisomerization is a smooth, monotonic function rising from zero at the instant of the stimulus to a positive value, before declining. Its time course is governed by a single time constant, k , so that the transmitter concentration is $y = f(kt)$, where t is the time elapsed after the light flash. (4) The time constant, k , is chosen so as to minimize the temporal width of the voltage generated for a given time to peak of the response t_p .

A direct consequence of assumptions (3) and (4) is that the transmitter concentration, $f(kt)$, increases almost logarithmically with respect to time in the neighbourhood of t_p . This is because, by assumption (4), k is chosen so as to maximize the rate of increase in transmitter concentration, $\partial f(kt)/\partial t$, at $t = t_p$; so that

$$\frac{\partial^2 f(kt)}{\partial k \partial t} = 0 \quad \text{at } t = t_p \quad (2)$$

However, from the identity

$$\frac{\partial f(kt)}{\partial \log t} = t \frac{\partial f(kt)}{\partial t} = k \frac{\partial f(kt)}{\partial k} \quad (3)$$

it follows that for our optimized system

$$\frac{\partial^2 f(kt)}{(\partial \log t)^2} = kt \frac{\partial^2 f(kt)}{\partial k \partial t} = 0 \quad \text{at } t = t_p \quad (4)$$

Integrating equation (4) twice with respect to $\log(t)$ yields $f(kt) = a \log(t) + b$ in the neighbourhood of t_p (a, b constant). The impulse-response function represents the sum of the voltages generated by all the channels that open following a light flash. This population of channels has, by assumption (2), thresholds to transmitter concentration that are normally distributed. If the concentration of transmitter rises logarithmically with time it follows that the temporal frequency of opening of channels will be log-normal so that, if channel opening times are brief, the voltage generated will follow equation (1). The model is illustrated in Fig. 2.

Our model describes the photoreceptor response as the result of a variable threshold associated with a process that is logarithmically related to time. Of the two temporal parameters only μ , which describes the rate of release of the internal transmitter, is significantly affected by temperature or light adaptation. We intend to continue our investigations by examining the response of 'faster' insect eyes, such as those of the fly and dragonfly. If equation (1) can be applied to other photoreceptors, then the two parameters, μ and σ , will provide a simple method of comparing their temporal frequency response, similar to the parameters describing the approximately gaussian angular sensitivity curves of photoreceptors¹⁷.

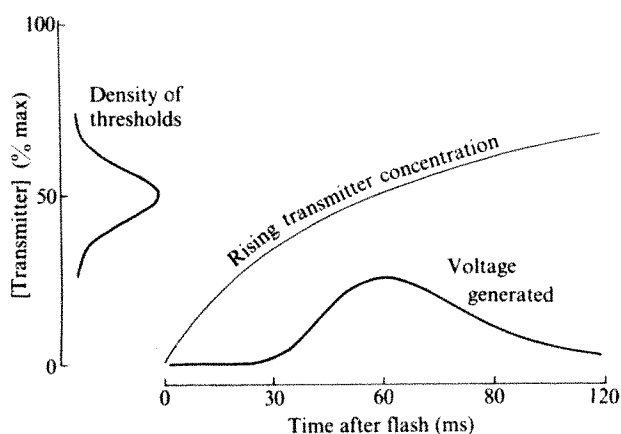


Fig. 2 Illustration of the physical model proposed in the text. The Gaussian curve describing the thresholds for the opening of channels is transformed into the asymmetric impulse-response through the non-linearity of the rising 'transmitter' concentration. The Langmuir isotherm, $y = kt/(1 + kt)$ is chosen to describe the transmitter concentration, y . In the region of the activation of channels, between $y = 0.3$ and 0.7 , this equation is approximated by $y = 0.25 \log(t) + c$. Choosing to activate channels in this logarithmic region provides for the least scatter in channel opening times.

Received 2 October 1980; accepted 15 January 1981.

1. Fuortes, M. G. F. & Hodgkin, A. L. *J. Physiol., Lond.* **172**, 239–263 (1964).
2. French, A. S. *Nature* **283**, 200–202 (1980).
3. Cone, R. A. in *Biochemistry and Physiology of Visual Pigments* (ed. Langer, H.) 275–282 (Springer, Berlin, 1973).
4. Hamdorf, K. & Kirschfeld, K. *Z. Naturforsch.* **35c**, 173–174 (1980).
5. Scholes, J. *Cold Spring Harb. Symp. quant. Biol.* **30**, 517–527 (1965).
6. Lillywhite, P. G. *J. comp. Physiol.* **122**, 189–200 (1977).
7. Pinter, R. B. *J. comp. Physiol.* **77**, 383–397 (1972).
8. Wilson, M. J. *J. comp. Physiol.* **97**, 323–328 (1975).
9. Payne, R. *Biophys. struct. Mech.* **6**, 235–251 (1980).
10. Baylor, D. A., Hodgkin, A. L. & Lamb, T. D. *J. Physiol., Lond.* **242**, 685–727 (1974).
11. Borsellino, A., Fuortes, M. G. F. & Smith, T. G. *Cold Spring Harb. Symp. quant. Biol.* **30**, 429–443 (1965).
12. Zettler, F. Z. *vergl. Physiol.* **64**, 432–449 (1969).
13. French, A. S. *Biol. Cybernet.* **32**, 115–123 (1979).
14. Gaddum, J. H. *Nature* **156**, 462–466 (1945).
15. Gaddum, J. H. *MRC spec. Rep.*, 183 (1933).
16. Bliss, C. I. *Ann. appl. Biol.* **24**, 134–166 (1937).
17. Washizu, Y., Burkhardt, D. & Streck, P. Z. *vergl. Physiol.* **48**, 413–428 (1964).

Choline biosynthesis by a preparation enriched in synaptosomes from rat brain

Jan K. Blusztajn & Richard J. Wurtman

Laboratory of Neuroendocrine Regulation, Department of Nutrition and Food Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

It is widely held that brain cells are unable to synthesize choline *de novo*, and that the only source of this compound for brain acetylcholine or membrane biosynthesis is the choline or choline-containing phospholipids taken up from the circulation¹. This notion has been difficult to reconcile with observations²⁻⁴ that there is a net efflux of choline from the brain. Recently we⁵ and others^{6,7} have demonstrated that various preparations of mammalian brain contain enzymes, the phosphatidylethanolamine *N*-methyltransferases (PeMT), which catalyse the synthesis of phosphatidylcholine (PC), using *S*-adenosylmethionine (SAM) as a methyl donor for the stepwise methylation of phosphatidylethanolamine (PE). The highest specific activity of PeMT was present in synaptosomal preparation^{5,6}. We now report that rat brain synaptosomal preparations can also metabolize the PC generated by PeMT to liberate free choline.

Male Sprague-Dawley rats (180–220 g, Charles River) were housed under light (Vita-Lite, Duro-Test) for 12 h daily and had free access to water and food (Charles River Rat Chow). Animals were decapitated, their brains quickly dissected and synaptosomal preparations of whole brains (excluding cerebellum) were obtained as described previously⁵; their identity was confirmed by electron microscopy. Synaptosomal preparations were incubated in Krebs-Ringer bicarbonate buffer (pH 7.4) at 37 °C. Heat-treated synaptosomes (100 °C, 20 s) were used to generate blank values. Phosphatidyl-*N,N*-dimethylethanolamine (P2) (Gibco) was used as PeMT substrate, and [*Me*-³H]*S*-adenosylmethionine (10 mCi μ mol⁻¹, NEN) as the methyl donor. Enzymatic reactions were stopped by addition of the appropriate extraction solution. Phospholipids were extracted and purified as described before⁵, and the ³H-methyl present in newly formed ³H-PC was quantified by liquid scintillation spectrophotometry. The choline, including [*Me*-³H]choline, was extracted from the incubation media with 0.5% tetraphenylboron (Aldrich) in heptanone (Aldrich), and back-extracted from the heptanone phase into 0.4 M hydrochloric acid. The aqueous layer was then dried under vacuum, and the dry residue redissolved in ethanol and subjected to

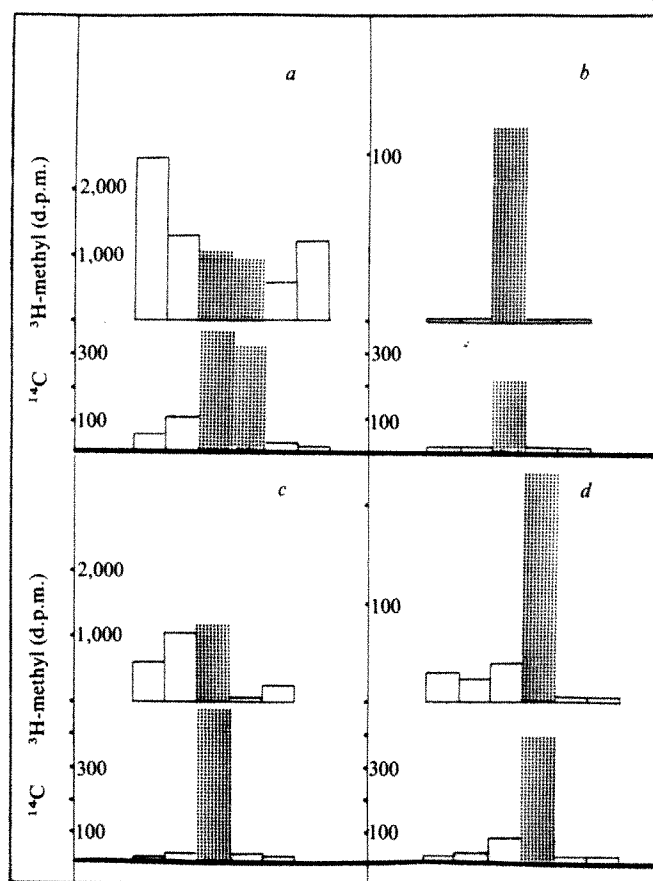


Fig. 1 Purification of newly synthesized choline. Synaptosomal preparations (0.5 mg protein) were incubated at 37 °C for 30 min in 0.1 ml Krebs-Ringer bicarbonate buffer (pH 7.4) containing: 4.5 μ M [*Me*-³H]SAM (10 mCi μ mol⁻¹), 50 μ g of ultrasound-dispersed P2, and 1,700 d.p.m. of [*Me*-¹⁴C]choline (60 mCi mmol⁻¹). The reaction was stopped by addition of 0.3 ml of 0.5% tetraphenylboron in heptanone. The choline in the heptanone was back-extracted to 0.3 ml of 0.4 M HCl, which was then dried under vacuum. The dry residue was then redissolved in 0.05 ml of ethanol and applied to cellulose TLC plates (a) or to filter paper for electrophoresis (c). The TLC solvent system contained: *n*-butanol/ethanol/acetic acid/water 40:10:1:15 (by volume). The electrophoretic conditions were: 0.75 M acetic acid, 0.75 M formic acid pH 2.0, constant voltage of 17 V cm⁻¹ for 40 min. Each sample also contained 5 μ g of choline chloride carrier. The plates and papers were stained for choline¹⁶, and segments staining positively were extracted with 50% ethanol in 0.2 M HCl. The extracts were then concentrated and rechromatographed as described above, except that TLC-derived material (a) was subjected to electrophoresis (b), while electrophoresis-derived material (c) was subjected to TLC (d). a, b, c and d show distributions of radioactivity on chromatograms prepared as described above. Shaded areas indicate segments which stained positively for choline. Lower panels show the distribution of ¹⁴C-choline standards and provide estimates of choline recovery. Incubations using heat-treated synaptosomal preparations yield flat background radioactivity patterns in b and d.

Table 1 Phosphorylation of newly synthesized choline by choline kinase

	Choline (d.p.m.)	Phosphorylcholine (d.p.m.)
Control	257 $\pm$ 22	0 (background)
Choline kinase	31 $\pm$ 7	323 $\pm$ 51

Incubations, extractions and TLC prepurification of newly synthesized choline were carried out as in Fig. 1 legend. Choline was extracted from TLC plates with 50% ethanol in 0.2 M HCl and the extracts were then dried under vacuum. The dry residues were resuspended in 0.1 ml of medium containing 12 mM glycylglycine, 8 mM dithiothreitol, 18 mM MgCl₂, 1.2 mM ATP and 0.05 U choline kinase (pH 7.4), and incubated for 15 min at 30 °C. Reaction was stopped with 0.1 ml ethanol and the mixtures concentrated and subjected to paper electrophoresis (as described in Fig. 1 legend). Segments were visualized for choline¹⁶ and phosphorylcholine¹⁷ and then the radioactivity counted. Data are expressed as means $\pm$ s.d.

paper electrophoresis⁸ and TLC. As described below, procedures were needed to reduce background radioactivity. The radioactivity in segments of the chromatograms was determined by liquid scintillation spectrophotometry. The recovery of choline was calculated by adding ¹⁴C-choline internal standard to synaptosomal incubations. The ³H-choline formed enzymatically was established by subjecting the water-soluble products of the methylation prepurified by TLC, to phosphorylation by choline kinase, purifying the products by paper electrophoresis. To determine the ratio of choline and PC in the incubations, a radioenzymatic assay of choline⁹, using choline [γ -³²P]ATP; the PC was purified by TLC and acidified before assay.

Identification of newly synthesized choline proved relatively difficult. It was not until we combined TLC and paper electrophoresis that we could see a sharp ^3H -methyl radioactive peak that co-migrated with authentic nonradioactive and ^{14}C -choline (Fig. 1). The heat-treated preparation of synaptosomes produced no radioactive product that would co-migrate with authentic choline in either of these chromatographic systems. After subjecting the ^3H -methyl material that had co-chromatographed with authentic choline to phosphorylation by choline kinase, virtually none of the radioactivity co-chromatographed with choline (Table 1). The radioactivity could, however, be recovered from material co-migrating with phosphorylcholine, as would be expected if the ^3H -methyl material were authentic choline.

The rate of incorporation of ^3H -methyl groups into choline depended on the concentration of PeMT substrate present—addition of increasing amounts of P2 to incubation mixtures (Fig. 2) caused parallel increases in the apparent rates of synthesis of PC and choline until saturation occurred at $\sim 500\ \mu\text{g}$ P2 per ml. This increase in apparent synthesis rate was more pronounced for ^3H -PC, where the ratio of rate of ^3H -methyl incorporation when 1 mg per ml of P2 was present to incorporation without added P2 was 5.7; the same ratio for ^3H -choline was only 1.5. Addition of dimethylethanolamine (30 mM final concentration) did not enhance the synaptosomal formation of choline (data not shown). These results indicate that the ^3H -choline must have been formed by degradation of newly synthesized PC, or, less likely, by action of some as yet undiscovered enzyme on a degradation product of P2.

The SAM concentration at which our assay was carried out ($4.5\ \mu\text{M}$) is below the apparent K_m for PeMT^{5,6}. However, at this SAM concentration the reaction proceeded linearly with time for up to 50 min. The incorporation of ^3H -methyl into ^3H -PC proceeded at the rate of 1.3 pmol per mg protein per 30 min, and into ^3H -choline at 0.4 pmol per mg protein per 30 min (Table 2). The choline newly formed during this interval would represent 88 p.p.m. of the total choline in the synaptosomal preparation, whereas the newly formed PC would represent only 1.9 p.p.m. of total PC (Table 2). Thus the relative enrichment of the choline pool was 47-fold greater than that of PC. These results suggest that there may exist specific physical domains of phospholipids within synaptosomal membranes, such that the PE methylation and PC degradation occur in the same domain. This would explain why up to 23% of the [^3H]choline present in the incubations could be recovered as free choline molecules and not covalently bound in the form of PC, while free choline constituted <1% of the total amount of

Table 2 Comparison of synthesis of phosphatidylcholine and choline by rat brain synaptosomal preparations

	Phosphatidylcholine	Choline
^3H -methyl incorporation (pmol per mg protein per 30 min)	1.323 ± 0.219	0.396 ± 0.086
Concentration (nmol per mg protein)	708 ± 29	4.5 ± 0.82
Specific activity (p.p.m.)	1.87 ± 0.37	88 ± 34
I_{Ch}	0.30	
I_{PC}		
SA_{Ch}	47.3	
SA_{PC}		

Incubation, extraction and purification conditions were as described in Fig. 1 and Table 1 legends. Free ^3H -choline and ^3H -choline hydrolysed off ^3H -PC were determined by radioenzymatic assay⁹. Data are expressed as means $\pm$ s.d. ^3H -methyl incorporations are I_{PC} and I_{Ch} , specific activities are SA_{PC} and SA_{Ch} for phosphatidylcholine and choline, respectively.

choline (that is free plus PC-bound) in the synaptosomal preparations (Table 2). Other studies have suggested the existence of two PC pools in synaptosomes, one with a $t_{1/2}$ of 2 days and another with $t_{1/2}$ of 52.5 days (ref. 10). Our results indicate the existence of yet another pool, which turns over much more rapidly, and which contains PC molecules derived from the PeMT pathway.

The brain is reported to contain several enzyme systems able to catalyse the degradation of PC to liberate free choline: phospholipases A_1 and A_2 and lysophospholipase¹¹ generate glycerol-3-phosphorylcholine, which can subsequently be hydrolysed in a one-step¹² or two-step process¹³ to yield free choline; phospholipase D can catalyse PC hydrolysis to liberate free choline directly¹⁴; lysophospholipase D could liberate choline from lysophosphatidylcholine formed by the action of phospholipase A_2 (ref. 15). Our data do not indicate which, if any, of these enzyme systems liberate ^3H -choline from ^3H -PC formed in synaptosomal preparation by the transmethylation pathway—our present techniques are insufficient to detect other [^3H]choline containing species, which might provide clues.

It seems clear that previously, by measuring only phospholipid products, we⁵ and others^{6,7} have underestimated the contribution of PeMT to brain choline. It is entirely possible that other compounds, such as glycerol-3-phosphorylcholine and phosphorylcholine, also accumulate in our experimental system as intermediates in the genesis of free choline; thus we may still be underestimating the total rate of choline production in the brain. The absolute contribution of *de novo* synthesis to total brain choline and the extent to which this source provides precursor molecules for acetylcholine formation await further study.

These studies were supported in part by a grant from the National Institute of Mental Health. We thank Mr Robert Scheinman for technical assistance.

Received 15 September 1980; accepted 20 January 1981.

1. Ansell, G. B. & Spanner, S. in *Cholinergic Mechanisms and Psychopharmacology* (ed. Jenden, D. J.) (Plenum, New York, 1977).
2. Aquilonius, S. M., Ceder, G., Lying-Tunell, U., Malmund, H. O. & Schubert, J. *Brain Res.* **99**, 422–430 (1975).
3. Choi, R. L., Freeman, J. J. & Jenden, D. J. *J. Neurochem.* **24**, 735–741 (1975).
4. Dross, K. & Kewitz, H. N.-S. *Archs Pharmac.* **275**, 91–106 (1972).
5. Blusztajn, J. K., Zeisel, S. H. & Wurtman, R. J. *Brain Res.* **179**, 319–327 (1979).
6. Crews, F. T., Hirata, F. & Axelrod, J. *J. Neurochem.* **34**, 1491–1498 (1980).
7. Mozzi, R. & Porcellati, G. *FEBS Lett.* **100**, 363–366 (1979).
8. Potter, L. T. & Murphy, W. *Biochem. Pharmac.* **16**, 1386–1388 (1967).
9. Goldberg, A. M. & McCaman, R. E. *J. Neurochem.* **20**, 1–8 (1973).
10. Pasquini, J. M., Krawiec, L. & Soto, E. F. *J. Neurochem.* **21**, 647–653 (1973).
11. Cooper, M. F. & Webster, G. R. *J. Neurochem.* **17**, 1543–1554 (1970).
12. Webster, G. R., Marples, E. A. & Thompson, R. H. S. *Biochem. J.* **65**, 374–377 (1957).
13. Abra, R. M. & Quinn, P. J. *Biochim. biophys. Acta* **380**, 436–441 (1975).
14. Saito, M. & Kanfer, J. *Archs Biochem. Biophys.* **169**, 318–323 (1975).
15. Wykle, R. L. & Schremmer, J. M. *J. biol. Chem.* **249**, 1742–1746 (1974).
16. Beiss, U. *J. Chromatography* **13**, 104–110 (1964).
17. Wade, H. E. & Morgan, D. M. *Nature* **171**, 529–530 (1953).

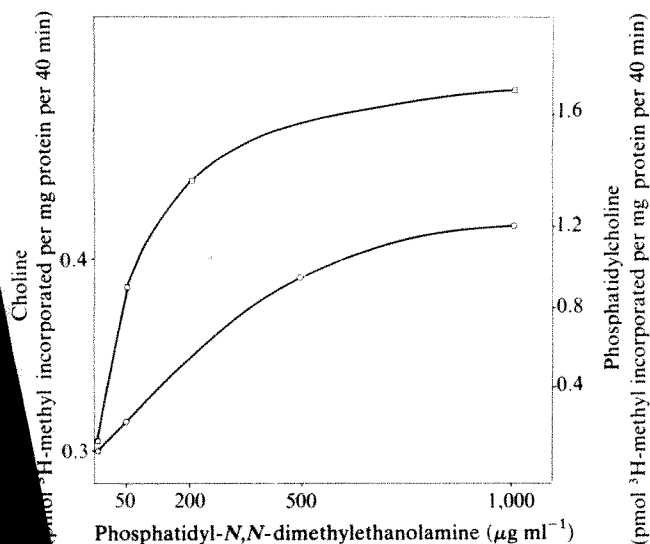


Fig. 2 Effects of P2 on PC (□) and choline (○) formation. Incubations were carried out for 40 min as described in Fig. 1 and Table 1 legend. Extractions and purification of products were performed as described in ref. 5 and Fig. 1 legend.

ATP-stimulated endoprotease is associated with the cell membrane of *E. coli*

Richard W. Voellmy* & Alfred L. Goldberg

Department of Physiology, Harvard Medical School, Boston, Massachusetts 02115, USA

Despite knowledge of the physiological significance and regulation of protein degradation in bacteria¹⁻⁷, the pathway of proteolysis and the responsible enzymes are still not known. Degradation of cell proteins in bacterial and animal cells requires continuous ATP production^{1,8}, inhibition of which in *Escherichia coli* prevents the degradation of normal proteins in growing cells⁸, accelerated breakdown of such proteins in starving cultures^{3,9} and the very rapid breakdown of abnormal proteins^{8,10}. Intracellular proteolysis proceeds by repeated endoproteolytic steps and ATP is required for the initial cleavages of the substrate¹⁰. We have recently demonstrated ATP stimulation of proteolysis in extracts of bacterial and animal cells¹¹⁻¹⁵. These ATP-stimulated systems seem to be responsible for the rapid degradation of abnormal proteins¹¹⁻¹⁴ *in vivo*, but they may also be involved in the catabolism of normal cell proteins^{1-3,9}, limited proteolysis, such as the processing of precursors for secreted or membrane proteins¹⁶⁻²⁰, and the selective inactivation of specific proteins, as occurs in the ATP-dependent cleavage of the lambda repressor by the recA protein²¹. We report here that membrane fragments contain an ATP-stimulated protease that degrades cell proteins to large peptides (of molecular weight (MW) 71,500) which are then rapidly hydrolysed to amino acids by soluble ATP-independent enzymes.

When extracts of freshly grown *E. coli* were prepared by gentle methods (by grinding with alumina, by using the French press at 3,000 p.s.i., or by freezing and thawing followed by lysozyme treatment and osmotic shock¹³), the degradation of [*Me*-¹⁴C]globin^{13,22} was stimulated two- to fourfold by addition of ATP (3 mM). With less gentle methods of lysis (the French press at 12,000–15,000 p.s.i., repeated freezing and thawing followed by lysozyme treatment, or sonication), the resulting extracts contained high levels of proteolytic activity but showed very little (<25%) stimulation by ATP¹³.

Approximately 70–80% of the proteolytic activity after gentle cell lysis was recovered on the pellet resulting from centrifugation at 160,000g for 60 min (Table 1). When these pellets were resuspended, they showed ATP-stimulated proteolysis. The proteolytic activity remaining in the supernatant was not affected by ATP (Table 1).

When these extracts were analysed by gel filtration on Sepharose 2B or 4B, most of the ATP-stimulated activity was eluted in the void volume ahead of 70S ribosomes (Fig. 1), and is thus associated with large particulate material. Some soluble proteolytic activity eluted with the bulk of cell protein of molecular weights less than that of β -galactosidase (MW 540,000). This latter activity (see below and refs 29, 39, 40) showed a much smaller stimulation by ATP than did the peak in the void volume, in which ATP increased proteolysis two- to fivefold. Identical elution patterns were obtained when gel filtration was repeated on the material in the void volume or was performed in the presence of 3 mM ATP. Therefore, ATP does not stimulate proteolysis by releasing the proteolytic activity from a bound to a free state.

This high-molecular-weight fraction contains polyribosomes and membrane fragments, as indicated by a high content of the (Ca²⁺ + Mg²⁺)ATPase (Fig. 2) and many membrane vesicles

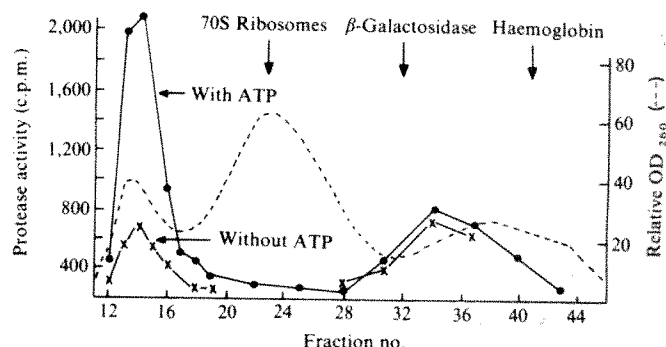


Fig. 1 Fractionation of *E. coli* proteolytic activity on Sepharose 4B. An extract from *E. coli*, prepared by grinding with alumina (50 mg protein), was filtered through an 80 × 2.5 cm column of Sepharose 4B equilibrated with 10 mM Tris-HCl, 10 mM MgCl₂, 50 mM KCl and 0.5 mM dithiothreitol (pH 7.5). Proteolytic activity was determined with (●) and without (×) an ATP-generating system as in Table 1 legend. β -galactosidase was assayed using *o*-nitrophenyl- β -D-galactoside as a substrate¹⁰. Similar data were obtained with extracts prepared as in Table 1 legend.

were evident in the electron microscope (in collaboration with Dr Richard Goldstein). When the extracts were incubated with puromycin (50 μ g ml⁻¹) for 5–10 min at 37 °C or with ribonuclease A (100 μ g ml⁻¹) before gel filtration to convert polysomes to monosomes²³, or when cells were incubated for 15 min at 10–15 °C before their collection to release ribosomes from the RNA²³, the ATP-stimulated proteolytic activity still eluted in the void volume of Sepharose 4B. Thus it could not have been associated with the polysomes.

Extracts subjected to sucrose density centrifugation showed maximal proteolytic activity at 47.5% sucrose (Fig. 2), the same density as membrane fragments prepared from marker cells grown in parallel in the presence of ³H-oleate to label membrane lipids. The maximal activity of the membrane-associated (Ca²⁺ + Mg²⁺)ATPase was also found in these same lipid-containing fractions (Fig. 2).

If the ATP-stimulated protease is membrane-associated, it should co-precipitate on incubation with an antibody directed against another membrane protein. As membrane vesicles prepared by grinding or French press are inverted²⁴, an antibody against an inner membrane protein, the F₁ ATPase, was used. When it was incubated with the fraction purified by gel filtration (Fig. 1), it is precipitated most (76–90%) of the membrane-bound ATPase (data not shown) and reduced the proteolytic activity in the supernatant (Table 2). By contrast, when the antibody was added to the soluble proteolytic activity obtained by gel filtration (Fig. 1), no inhibition of globin hydrolysis was observed. A control immunoprecipitation was also carried out with purified exogenous elongation factor G (EFG) and an antibody directed against this protein (Table 2). The concentrations of EFG and anti-EFG were adjusted such that a similar amount of protein precipitated in these conditions to that which precipitated with the antibody against the membrane ATPase. Virtually none of the membrane-associated proteolytic activity was lost from the supernatant.

Previous studies^{1,25,26} have demonstrated appreciable proteolytic activity in membrane fractions prepared by low speed centrifugation. We therefore tested whether this fraction contained the ATP-stimulated proteolytic system. Cells lysed by freezing and thawing followed by lysozyme treatment and osmotic shock, and the extracts were centrifuged for 10–15 min to remove unbroken cells and then at 25,000g for 20 min. The pellets of the second spin were very rich in hydrolysing ¹⁴C-globin, but this reaction was not stimulated by ATP (3 mM). However, in the presence of 10 mM sodium cyanide (an ATPase inhibitor), ATP caused a clear stimulation of globin degradation (data not shown), suggesting that a (Ca²⁺ + Mg²⁺)ATPase is not essential for the stimulation of proteolytic activity, in accord with findings on

* Present address: Department of Molecular Biology, University of Geneva, Geneva, Switzerland.

Table 1 Distribution of the ATP-stimulated proteolytic activity on ultracentrifugation

Fractions	Proteolytic activity (c.p.m.)		% Stimulation by ATP
	Without ATP	With ATP	
Crude extract	5,490	14,260	160
Pellet after second wash	3,831	8,772	129
Supernatant of first wash	1,850	1,860	—
Supernatant of second wash	804	792	—

E. coli K12 strain NF172 (ref. 9, 500-ml portions) was grown in Luria broth⁴² at 37 °C with vigorous aeration. The cells were cooled rapidly and collected at an A_{550} of 1.0. Three litres of cells were lysed by grinding with alumina for a short period. The resulting crude extract in buffer I (25 mM Tris-HCl pH 7.5, 20 mM MgCl₂, 0.5 mM dithiothreitol) was then centrifuged at 12,000g for 20 min. The supernatant was dialysed overnight against buffer I, and an aliquot of this crude extract was assayed for proteolytic activity. The dialysed extract was spun at 120,000g for 75 min. The pellets were resuspended in this buffer and centrifuged again. The pellet from this first wash was resuspended once more and centrifuged as before. The final pellets were resuspended in the original volume of buffer I, and together with the supernatants of the washes were assayed for proteolytic activity in the presence or absence of ATP and an ATP-generating system. The hydrolysis of [*Me*-¹⁴C]- (apo)haemoglobin^{13,22} to acid-soluble material was used to estimate proteolytic activity. Incubation mixtures contained 1 ml of extract, 10 μ l of ¹⁴C-globin (~15,000 c.p.m.) and 50 μ l of an ATP solution (60 mM) or an ATP-generating system (60 mM ATP, 250 mM phosphocreatine, 0.4 mg ml⁻¹ creatine phosphokinase, rabbit muscle type I [Sigma]). Incubations were carried out at 30 °C for 60–90 min. Comparisons between the absolute proteolytic activity in the crude extracts, supernatants and pellets may be misleading as these fractions contained very different amounts of nonradioactive proteins that may compete with the ¹⁴C-labelled substrate. Qualitatively, similar data were obtained when the pellets were resuspended in buffer I containing 0.5 M KCl or 1 mM Tris-HCl (pH 7.5), although these conditions reduced the degree of stimulation by ATP.

Furthermore, in related experiments (data not shown) we have demonstrated ATP-stimulated proteolysis in membrane vesicles of mutants²⁷ that lack a functioning (Ca²⁺ + Mg²⁺) ATPase (AN120, an *uncA*⁻ strain).

The acid-soluble products of protein breakdown were analysed by gel filtration on Sephadex G-15. Crude *E. coli* extracts, prepared by alumina grinding, and the membrane fraction, isolated by gel filtration on Sepharose 4B, were incubated in parallel with either ¹⁴C-globin or ³⁵S-labelled *E. coli* protein (denatured with acid to inactivate proteases). Both substrates were hydrolysed on these extracts to acid-soluble material by an ATP-stimulated process (stimulation was 100–200% with globin and 70–100% with the *E. coli* proteins). The crude extracts degraded the ³⁵S-proteins to amino acids (Fig. 3) and thus can carry out the complete degradation of cell proteins, in similar fashion to intact cells. The [*Me*-¹⁴C]globin was hydrolysed to dipeptides. Related observations suggest that the labelled methyl group on the lysine prevents the hydrolysis of dipeptides by certain amino-peptidases.

When either substrate was incubated with the purified membrane fraction, most of the resulting acid-soluble ³⁵S- or ¹⁴C-products eluted close to the void volume. Very little was in the form of amino acids or even dipeptides (Fig. 3). The sizes of these proteolytic products did not differ whether or not ATP was present (data not shown). The ATP-stimulated proteolytic activity generated mainly peptides of MW 1,500 or greater²⁸ which were not degraded further. Thus the rate-limiting step in protein degradation in these extracts is an ATP-stimulated endoproteolytic cleavage, in accord with *in vivo* findings¹⁰. This initial step is membrane associated, while soluble cytoplasmic enzymes (exo- and endopeptidases^{6,29,30}) catalyse the subsequent conversion of the peptides to free amino acids. Thus, the ³⁵S-labelled peptides generated by the membrane fraction were mixed with the soluble fraction obtained by

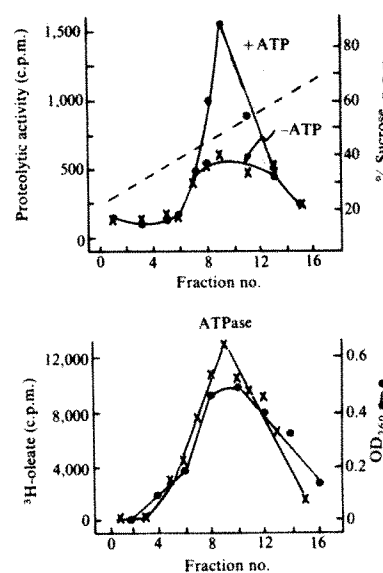


Fig. 2 Sucrose-density centrifugation of the particulate proteolytic activity. An extract prepared with the French press at 3,000 p.s.i. was centrifuged at 12,000g for 20 min. The supernatant was spun for 60 min at 12,000g and the resulting pellet washed once and resuspended in buffer I. For measurements of proteolytic activity against ¹⁴C-globin (see Table 1) and ATPase activity, 5–10 mg of the pelleted material were applied to linear sucrose gradients (32.5 ml) ranging from 25–65% and spun for 24 h at 25,000 r.p.m. in a SW27 rotor. Fractions (65 drops) dialysed against buffer I to remove the sucrose were assayed for proteolytic activity. To identify the fractions containing membrane lipids, cells were grown in medium containing ³H-oleate (28.5 mCi per mg, NEN). Extracts from the ³H-oleate-labelled cells were spun in parallel, and the radioactivity in each fraction measured (lower panel). To assay ATPase activity, 0.5-ml portions of the fractions were incubated with 0.5 mM ATP for 45 min at 30 °C. The samples were then boiled, and their ATP content measured using the firefly luciferase assay^{8,9}.

centrifugation (Table 1) of the crude extract, they were quickly hydrolysed to free amino acids. This latter reaction was not stimulated by ATP.

Although a large fraction, perhaps most, of the ATP-stimulated proteolytic activity is associated with the cell membrane, we cannot conclude what per cent of the proteolytic activity is actually membrane bound *in vivo* or even in these cell-free extracts, because the soluble and membrane fractions contain very different amounts of peptidases (Fig. 3) and of nonradioactive proteins that serve as competing substrates. A localization of the ATP-stimulated proteolysis on the inner cell membrane seems likely as alumina grinding and the French Press at low pressures create primarily inverted membrane vesicles, which should be impermeable to ATP and ¹⁴C-globin. Furthermore, protein breakdown in intact cells is completely dependent on the continuous production of ATP^{1,3,8–10} and therefore requires contact with the intracellular nucleotide pools. However, these studies have not eliminated conclusively a possible association with the outer cell membrane³¹. As the existence of proteases free within a cell might be harmful, an association with the membrane could reduce the frequency at which critical enzymes are inappropriately hydrolysed. This localization may also affect the selectivity of protein breakdown. Bacteria, like animal cells, degrade soluble proteins with highly abnormal conformations especially rapidly^{1,2,4–7,10–12}. Such partially unfolded proteins probably have more hydrophobic regions exposed on their surface than do normal ones^{1,12} and, therefore, may bind differentially to the cell membrane. It is perhaps relevant that when bacteria make large amounts of abnormal proteins, these molecules associate in intracellular aggregates that lie adjacent to the surface membrane^{32,33}.

On the cell membrane, the ATP-stimulated endoprotease may also be important in the turnover of membrane components

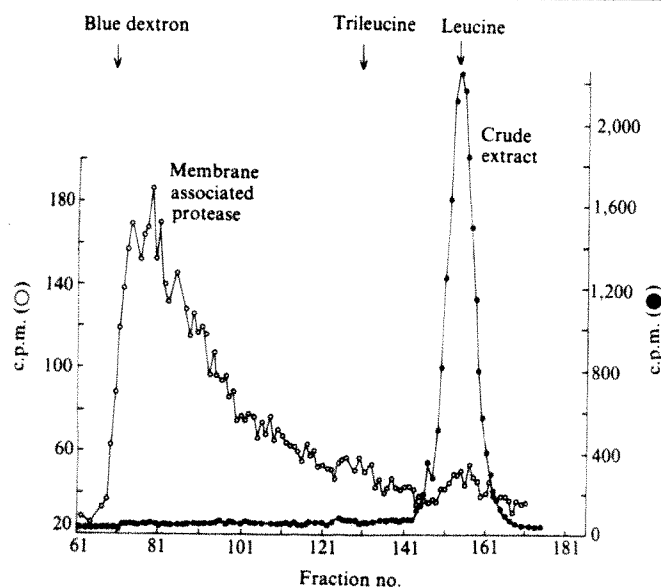


Fig. 3 Size distribution of the acid-soluble material produced by the membrane-associated protease and the crude extract. Cells were lysed by grinding with alumina. One portion of the extract was filtered through Sepharose 4B, and the fractions containing the membrane-associated protease were pooled and concentrated by ultrafiltration. This material (○) and a portion of the crude extract (●) containing an equal amount of protein were incubated with ^{35}S -labelled denatured *E. coli* protein for 2 h at 30 °C in the presence of an ATP-generating system. The protein was then precipitated with 10% trichloroacetic acid, and the supernatants applied to a 2 × 78 cm column of Sephadex G-15 in 0.05 M Tris-HCl (pH 7.5). Trileucine and leucine were run as molecular weight markers and were assayed with fluorescamine²⁶. To prepare ^{35}S -labelled *E. coli* proteins, 10 ml of strain NF 161 (*spoT⁻, arg⁻, met⁻*) were grown on glucose minimal medium⁸ containing arginine (60 $\mu\text{g ml}^{-1}$), a limiting amount of methionine (1.5 $\mu\text{g ml}^{-1}$) and 0.5 mCi of ^{35}S -methionine (460 Ci nmol⁻¹, NEN). When growth ceased, the cells were centrifuged and washed with 10 ml of growth medium lacking ^{35}S -methionine. The cells were then treated with 10% trichloroacetic acid. After centrifugation, the pellet was dissolved in 1 M NaOH and the pH adjusted to 7.0 with HCl. Before use, the suspension was dialysed extensively against buffer I.

and in the processing of secreted periplasmic or membrane proteins from larger precursors^{16-19,21} a process believed to occur on the inner cell membrane^{17,21,37}. Inner membrane vesicles are able to process membrane protein^{17,20}, and this reaction seems to depend on the high-energy membrane state^{34,35}. Further characterization of the 'signal peptidase(s)' and the membrane-associated globin-degrading activities should clarify whether these proteolytic systems are in some way related. In addition, the ATP-stimulated enzyme may be involved in the rapid degradation of the released signal peptides.

These findings raise the possibility that the effect of ATP is mediated by some membrane-related function. However, there is strong evidence that the high-energy membrane state is not required for protein breakdown *in vivo*⁸, and uncoupling agents do not inhibit the ATP effect in these membrane preparations (data not shown). The present studies have also ruled out an involvement of the $(\text{Ca}^{2+} + \text{Mg}^{2+})\text{ATPase}$. Moreover, when *E. coli* are lysed by less gentle methods than were used here, we can isolate from the soluble fraction an ATP-dependent endoprotease, as well as seven ATP-independent proteases^{29,38-40}. Therefore, a membrane-association is not essential for the effect of ATP on protein degradation. After purification^{29,41}, the soluble ATP-dependent enzyme resembles strongly the membrane-associated activity described here²⁹. Possibly the soluble enzyme is derived from the membrane fraction, or alternatively the bacteria may contain two distinct ATP-stimulated proteolytic systems that serve different physiological functions.

Table 2 Loss of proteolytic activity on treatment with antibody against membrane-associated ATPase

Source of proteolytic activity	Additions	Activity remaining in supernatant (c.p.m.)	% Proteolytic activity lost from supernatant
Particulate fraction	Rabbit IgG	4,661	0
	anti-ATPase (EFG) + anti-EFG	2,876	39
Soluble fraction	Rabbit IgG	4,455	4
	Anti-ATPase	5,677	0
		5,239	8

An extract prepared by grinding with alumina was chromatographed through Sepharose 4B. The peaks of proteolytic activity in the void volume (particulate activity) and in the soluble fraction were separated by gel chromatography as in Fig. 1. Portions (1 ml) of the two fractions in buffer I were incubated with antibody for 24 h at 4 °C. Precipitates were removed in a refrigerated clinical centrifuge, and proteolytic activity against ^{14}C -globin in the supernatants was assayed in the presence of an ATP-generating system (Table 1). The purified IgG fractions ($A_{280} = 0.57$) or purified EFG ($A_{280} = 0.02$) were added to 1 ml of the protease preparations. All antisera were purified by ammonium sulphate fractionation and chromatography on diethylaminoethyl and CM-cellulose⁴³ to remove the associated proteolytic activity. The purified IgG fractions contained no measurable activity against ^{14}C -globin.

We thank Dr P. C. Tai for advice and for providing pure ribosomes and ribosomal subunits, Dr F. Gibson for strain AN120, Professor L. Heppel for the antibody against membrane ATPase, and Ms Robin Levine for assisting in the preparation of this manuscript. During this work, R. W. V. was a Fellow of the Swiss National Fund. These studies were supported by grants from the National Institute of Neurobiological Diseases and Stroke and the Juvenile Diabetes Fund.

Received 15 May; accepted 12 November 1980.

- Goldberg, A. L. & St John, A. C. *Rev. Biochem.* **45**, 747-803 (1976).
- Pine, M. A. *Rev. Microbiol.* **26**, 103-126 (1972).
- Mandelstam, J. *Bact. Rev.* **24**, 289-308 (1960).
- Zipser, D. & Bhausaar, P. *J. Bact.* **127**, 1538-1542 (1976).
- Miller, C. G. in *Limited Proteolysis in Microorganisms* (eds Holtzer, H. & Cohen, G.) 65-69 (US Dept of Health, Education & Welfare, DHEW Publ. no. 79-1591, 1979).
- Olden, K. & Goldberg, A. L. *Biochem. biophys. Acta* **542**, 385-398 (1978).
- St John, A. C. & Goldberg, A. L. *J. biol. Chem.* **253**, 2705-2711 (1978).
- Kowit, J. D. & Goldberg, A. L. *J. biol. Chem.* **252**, 8350-8357 (1977).
- Etlinger, J. D. & Goldberg, A. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 54-58 (1977).
- Goldberg, A. L., Kowit, J. D. & Etlinger, J. D. & Klemes, Y. in *Protein Turnover and Lysosomal Function* (eds Segal, H. & Doyle, D.) 171-196 (Academic, New York, 1978).
- Murakami, K., Voellmy, R. & Goldberg, A. L. *J. biol. Chem.* **254**, 8194-8200 (1979).
- Voellmy, R., Murakami, K. & Goldberg, A. L. in *Limited Proteolysis in Microorganisms* (eds Holtzer, H. & Cohen, G.) 7-17 (US Dept of Health, Education & Welfare, DHEW Publ. no. 79-1591, 1979).
- DeMartino, G. N. & Goldberg, A. L. *J. biol. Chem.* **254**, 3712 (1979).
- Blobel, G. & Dobberstein, B. *J. Cell. Biol.* **67**, 852-862 (1975).
- Chang, C. N., Blobel, G. & Model, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 361-365 (1978).
- Inouye, H. & Beckwith, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1440-1444 (1977).
- Smith, W. P., Tai, P. L., Thompson, R. C. & Davis, B. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2830-2834 (1977).
- Chang, C. N., Inouye, H., Model, P. & Beckwith, J. *J. Bact.* **142**, 726-728 (1980).
- Roberts, J. W., Roberts, C. W. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2283-2287 (1977).
- Etlinger, J. D. & Goldberg, A. L. *J. biol. Chem.* **255**, 4563-4568 (1980).
- Subramanian, A. R., Davis, B. D. & Beller, R. J. in *Cold Spring Harb. Symp. quant. Biol.* **34**, 223-225 (1969).
- Futa, M. *J. Membrane Biol.* **15**, 15-28 (1974).
- Regnier, P. & Thang, M. N. *FEBS Lett.* **36**, 31-33 (1973).
- Kowit, J., Choy, W. N., Champe, S. P. & Goldberg, A. L. *J. Bact.* **128**, 776-784 (1976).
- Butlin, J. D., Cox, G. B. & Gibson, F. *Biochem. J.* **124**, 75-81 (1971).
- Reiland, J. *Meth. Enzym.* **22**, 287-321 (1971).
- Voellmy, R., Rosenthal, E., Gronastajski, R. & Goldberg, A. L. *Cell* (in the press).
- Miller, C. G. & Schwartz, G. *J. Bact.* **135**, 603-611 (1978).
- Osborn, M. J., Gander, J. E., Parisi, E. & Carson, T. *J. biol. Chem.* **247**, 3962-3972 (1972).
- Prouty, W. F. & Goldberg, A. L. *Nature new Biol.* **240**, 147-150 (1972).
- Prouty, W. F., Karnovsky, M. J. & Goldberg, A. L. *J. biol. Chem.* **250**, 1112-1122 (1975).
- Ito, K., Mandel, G. & Wickner, W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1199-1203 (1979).
- Date, T., Zwizinski, C., Ludmerer, S. & Wickner, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 827-831 (1980).
- Nelson, N. & Schatz, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4365-4369 (1979).
- Zwizinski, C. & Wickner, W. *J. biol. Chem.* (in the press).
- Swamy, K. H. S. & Goldberg, A. L. (submitted).
- Goldberg, A. L., Swamy, K. H. S., Chung, C. H. & Larimore, F. S. in *Meth. Enzym.* (in the press).
- Goldberg, A. L., Strnad, N. P. & Swamy, K. H. S. *Ciba Fdn Symp.* **75**, 227-251 (1980).
- Larimore, F., Waxman, L. & Goldberg, A. L. (submitted).
- Miller, J. *Experiments in Molecular Genetics*, 433 (Cold Spring Harbor Laboratory, New York, 1972).
- Palacios, R., Palmiter, R. D. & Schimke, R. T. *J. biol. Chem.* **247**, 2316-2321 (1972).

Two mutations that alter the regulatory activity of *E. coli* *recA* protein

Jeffrey W. Roberts & Christine W. Roberts

Section of Biochemistry, Molecular and Cell Biology,
Wing Hall, Cornell University, Ithaca, New York 14853, USA

The *Escherichia coli* *recA* gene product is both a direct participant and a central regulatory element in important processes of DNA repair, including one that is probably responsible for all radiation mutagenesis and some chemical mutagenesis. It has the direct function of catalysing pairing of single-stranded DNA to a homologous region in duplex DNA^{1,2}, a reaction thought to be fundamental to genetic recombination. This activity of *recA* protein probably contributes to DNA repair by promoting recombination between damaged DNA molecules³. We have shown that *recA* protein also has a regulatory function, mediated by its ability to destroy repressors and possibly other proteins by proteolytic cleavage, leading to the induction of certain genes⁴⁻⁶; these include *recA* itself, regulated by the *lexA* gene product⁷⁻⁹, other genes probably involved in DNA repair, some of which are also regulated by the *lexA* gene product¹⁰, and the early genes of temperate bacteriophages, which are regulated by their repressors. *recA* protein is activated to cleave repressors *in vitro* by interaction with ATP (or the ATP analogue ATP- γ S) and a polynucleotide such as single-stranded DNA⁴, probably the same interaction that initiates the strand-pairing reaction¹¹. We have proposed that *recA* protein is also activated to attack repressors *in vivo* when it interacts with single-stranded DNA; DNA-damaging treatments such as UV-ray irradiation would thereby invoke the *recA*-dependent functions of recombination, repair, mutagenesis and prophage induction⁴. As further evidence that the proteolytic activity of *recA* protein is responsible for its regulatory function, we show here that the ability of two mutationally altered *recA* proteins to cleave phage λ repressor correlates with the ability of the mutant cells to induce prophage.

Both mutations *recA142* and *recA430* (formerly named *lexB30*) prevent phage repressor inactivation after inducing treatments, but they affect *recA* function in the cell differently. (1) The mutant *recA142* is defective in genetic recombination, like classical *RecA*⁻ mutants, and it does not induce λ prophage after normal inducing treatments¹². Unlike most *RecA*⁻ mutations, however, *recA142* supports spontaneous induction, an unprovoked but *recA*-dependent basal level of induction that occurs in wild-type lysogens (ref. 12 and Table 1). A λ repressor fragment characteristic of *recA*-mediated cleavage results from this spontaneous induction in both wild-type and *recA142* lysogens¹³. We thus expected that *recA142* protein would be able to cleave λ repressor in some conditions *in vitro*, but might have some defect in the interaction with DNA thought to initiate both strand pairing and repressor cleavage. (2) The mutant *recA430* was selected by Devoret and Blanco^{14,15} to be defective in λ induction, and it neither responds to inducing agents nor displays spontaneous induction (ref. 14 and Table 1). Furthermore, *recA* protein is not amplified in a *recA430* mutant cell in response to inducing treatments^{16,17}. Although the *recA430* mutation maps in the *recA* gene^{17,18}, conjugation crosses in the *recA430* cell produce nearly normal numbers of recombinant progeny. We expected that *recA430* protein would be unable to cleave λ repressor, but might be normal in activities required for genetic recombination.

To provide a source of these mutant proteins, we used strains carrying the *lexA* gene mutation *spr-51* (ref. 7), which is thought to inactivate a repressor of the *recA* gene and thus cause overproduction of *recA* protein^{5,19}. We confirmed that these derivatives maintain the *recA* phenotypes of the original

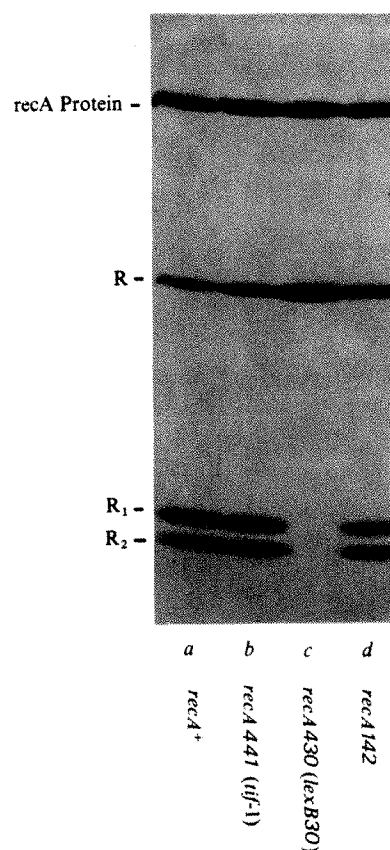


Fig. 1 Repressor cleavage directed by purified mutationally altered *recA* proteins. Reaction mixtures of 30 μ l were constructed as described previously⁴, except that they contained 2 mM MgCl₂, 1 mM ATP- γ S, 0.047 M NaCl, 0.013 M NaCl, 1.5 μ g *recA* protein, 3 μ g λ repressor and 0.039 μ g denatured λ DNA. After incubation for 3 h 20 min at 37°C, the reaction mixtures were subjected to SDS-gel electrophoresis and the gel was stained as described previously⁴. The *recA* protein concentration was calculated from a measured absorbance at 280 nm of 0.60 units mg⁻¹ ml⁻¹, corrected from the previously assumed 1.0 units mg⁻¹ ml⁻¹ (ref. 4). Track a, *recA*⁺ protein; track b, *recA441* protein (formerly named *tif-1*); track c, *recA430* protein; track d, *recA142* protein. R, λ repressor; R₁ and R₂, proteolytic fragments of λ repressor. *recA* protein and λ repressor were prepared as described elsewhere⁴. *recA430* protein was prepared from the *spr-51 recA430* strain KM1193, obtained from K. McEntee. *recA142* protein was prepared from the *spr-51 recA142* strain JR235, constructed by P1 transduction of DM1420 (ref. 7) to *srl300*:Tn10, followed by transduction to *Sr1*⁺ by P1 grown on the *recA142* strain JC4728 and screening for sensitivity to UV-light irradiation.

mutants, in that a λ lysogen of *spr-51 recA142* displays spontaneous induction whereas a lysogen of *spr-51 recA430* does not (data not shown). Like its *spr*⁺*recA142* parent, the *spr-51 recA142* strain shows by patch tests the high sensitivity to killing by UV-ray irradiation characteristic of *recA*⁻ mutants (data not shown). The parental *spr*⁺*recA430* mutant is moderately UV sensitive¹⁵; its *spr-51* derivative is less sensitive by patch tests, but still distinctly more sensitive than wild type.

We tested the ability of *recA* protein purified from *recA142* and *recA430* bacteria to cleave λ repressor in the single-stranded DNA and ATP- γ S dependent reaction. Figure 1 shows that *recA142* protein is active, whereas *recA430* protein is not. In control reactions, both wild-type *recA*⁺ protein and protein from the *recA* mutant *tif-1* (refs 5, 20) (renamed *recA441*), which displays constitutive expression of *recA*-dependent functions, are active. From densitometric scanning of stained repressor fragments produced during a time course, we determined that the *recA142* protein cleaves λ repressor at ~75%

the rate of wild type in the presence of ATP- γ S (11 compared with 15 ng repressor cleaved per min per μ g recA protein). In contrast, Fig. 1 implies that recA430 protein could cleave λ repressor at no more than 2% the rate of wild type. We tested the mutants in conditions that provide the highest cleavage rate with wild-type recA protein⁴, but observed no repressor cleavage by recA430 protein in other conditions—ATP instead of ATP- γ S, higher concentrations of DNA, or oligonucleotides substituted for λ DNA.

recA protein is a single-stranded DNA-dependent ATPase^{4,21-23} whose activity is probably essential to strand exchange, as this requires ATP hydrolysis^{2,11}. Figure 2 shows that recA430 protein has normal ATPase activity, consistent with its normal function in recombination. This result also shows that the protein we prepared was not destroyed during cell lysis and purification. The recA430 protein also is equivalent to the wild type in DNA-dependent entrapment of radioactive ATP- γ S on a membrane filter (data not shown), and is active in reannealing homologous single strands of DNA, possibly a partial reaction of the strand-pairing activity⁹. Thus, we conclude that recA430 protein is specifically defective in cleaving phage λ repressor, and that this defect accounts for the failure of λ to be induced in recA430 bacteria.

The recA142 protein, which does not effectively support recombination *in vivo*, is also an active ATPase *in vitro* (Fig. 2). However, its activity is much more sensitive to ionic strength than that of wild-type recA protein, displaying only 10% of the activity of the wild type at 0.20 M NaCl. This suggests a weaker interaction of the recA142 protein with DNA in the presence of ATP that could account both for the recombination deficiency of the mutant and for its failure to cleave λ repressor in inducing conditions: it may be inefficient in forming the ternary complex with ATP and DNA postulated to be the active intermediate in both strand pairing and repressor cleavage. In support of this interpretation, we have found that ATP, in contrast to the analogue ATP- γ S, promotes λ repressor cleavage much less efficiently with recA142 protein than with wild-type recA⁺ protein (<0.2 compared with 2 ng repressor cleaved per min per μ g recA protein). We presume that recA142 protein shows substantial cleavage activity with ATP- γ S because this analogue, unlike ATP, interacts essentially irreversibly with recA protein (unpublished experiments) and thus overcomes the reluctance of recA142 protein to form the active ternary complex.

The ATPase activity of the recA430 protein is only slightly more sensitive to ionic strength than the wild type (Fig. 2), presumably corresponding to its nearly normal activity in genetic recombination.

These experiments provide further direct evidence that cleavage of repressors by recA protein is the mechanism of repressor inactivation in induction. Because the recA430 protein supports recombination but not λ repressor inactivation *in vivo*, induction cannot result from the activity of recA protein as

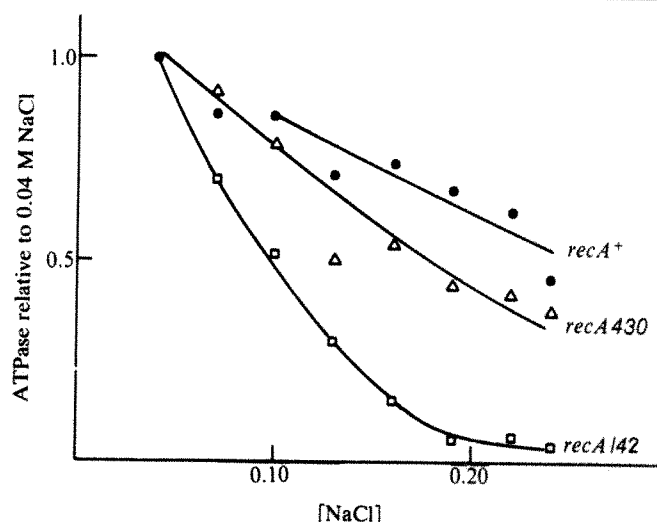


Fig. 2 Sensitivity of ATPase activity of mutationally altered recA proteins to ionic strength. The DNA-dependent ATPase activity was measured as described previously⁴, except that the concentration of recA protein was 100 μ g ml⁻¹ and the time of incubation was 30 min. Activities measured at 0.04 M NaCl were: recA⁺, 0.31 μ mol min⁻¹ mg⁻¹; recA430, 0.30 μ mol min⁻¹ mg⁻¹; recA142, 0.39 μ mol min⁻¹ mg⁻¹.

a recombination enzyme. Our results argue against models that postulate that the interaction of λ repressor with DNA (or a recA-DNA complex) is sufficient to promote inactivation of λ repressor in the absence of cleavage²⁴.

Besides their defect in λ repressor inactivation, recA430 bacteria also fail to carry out other recA-dependent processes, such as amplification of recA protein and radiation mutagenesis^{17,25}. We suggest that these defects arise at least in part from the inability of recA430 protein to inactivate the *E. coli* lexA protein, which is thought to be a repressor of the recA gene. The lexA protein is cleaved by recA protein⁸ in the conditions we have described for λ repressor inactivation; this supports the model that recA protein amplification occurs when lexA protein is cleaved in inducing conditions by recA protein present at a low basal level. Failure of recA430 protein to cleave lexA protein would thus prevent amplification of recA protein. The inability of the recA430 cell to be mutagenized by UV light could result from its failure to induce genes required for UV mutagenesis that are repressed by lexA or are regulated by other proteins sensitive to cleavage by recA protein. This is supported by the discovery of Kenyon and Walker¹⁰ that numerous *E. coli* genes involved in DNA repair are repressed by lexA and require recA function to be induced; these might include genes required for UV mutagenesis. Although the possibility has not been tested, it is likely that a third recA-dependent activity, lethal filamentation, is defective in recA430, as this process is also thought to result from inactivation of lexA protein by recA protein^{9,26}.

We thank N. L. Craig and E. M. Phizicky for advice and discussions, N. L. Craig for performing some preliminary experiments, A. J. Clark, R. Devoret, D. Botstein, D. Mount and K. McEntee for bacterial strains, and V. Vogt for criticizing the manuscript. This work was supported by NIH grant GM21941 and by a NIH Career Development Award to J.W.R.

Received 7 October 1980; accepted 29 January 1981.

Table 1 Spontaneous and mitomycin C induction of λ in recA lysogens

Strain	No. of infectious centres per ml of culture	
	Spontaneous	Mitomycin C treated
recA ⁺ (λ ⁺)	1 $\times$ 10 ⁶	1 $\times$ 10 ⁹
recA12(λ ⁺)	<10 ¹	<10 ¹
recA142(λ ⁺)	5 $\times$ 10 ⁵	7 $\times$ 10 ⁴
recA430(λ ⁺)	2 $\times$ 10 ² (clear)	2 $\times$ 10 ² (clear)

Cultures were grown in minimal medium¹³ at 37 °C to 5 $\times$ 10⁷ cells per ml. Portions of each were grown a further 90 min in the presence and absence of 5 μ g ml⁻¹ mitomycin C, treated with chloroform and titred. Phage from the recA430(λ ⁺) culture made clear plaques, indicating that they were derived not from induction but from growth following mutational alteration of the repressor. Bacterial strains have been described elsewhere¹³, except the recA430 strain JM776, which was obtained from R. Devoret. recA12 is a standard RecA⁻ mutation.

- Shibata, T., DasGupta, C., Cunningham, R. & Radding, C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1638-1642 (1979).
- McEntee, K., Weinstock, G. & Lehman, I. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2615-2519 (1979).
- Rupp, W. & Howard-Flanders, P. *J. molec. Biol.* **31**, 291-304 (1968).
- Craig, N. L. & Roberts, J. W. *Nature* **283**, 26-30 (1980).
- Roberts, J., Roberts, C. & Craig, N. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4714-4718 (1978).
- Phizicky, E. M. & Roberts, J. W. *J. molec. Biol.* **139**, 319-328 (1980).
- Mount, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 300-304 (1977).

8. Little, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 3225–3229 (1980).
9. Brent, R. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1932–1936 (1980).
10. Kenyon, C. & Walker, G. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2918–2923 (1980).
11. Cunningham, R. P., Shibata, T., DasGupta, C. & Radding, C. M. *Nature* **281**, 191–195 (1979).
12. Clark, A. J. *Rev. Genet.* **7**, 67–86 (1973).
13. Roberts, J. W. & Roberts, C. W. *Proc. natn. Acad. Sci. U.S.A.* **72**, 147–151 (1975).
14. Devoret, R. & Blanco, M. *Molec. gen. Genet.* **107**, 272–280 (1970).
15. Morand, P., Blanco, M. & Devoret, R. *J. Bact.* **131**, 572–582 (1977).
16. West, S. C. & Emmerson, P. *Molec. gen. Genet.* **151**, 51–67 (1977).
17. McEntee, K. *ICN-UCLA Symp.* **9**, 349–360 (1978).
18. Morand, P., Goze, A. & Devoret, R. *Molec. gen. Genet.* **157**, 69–82 (1977).
19. Gudas, L. & Mount, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5280–5284 (1977).
20. Castellazzi, M., George, J. & Buttin, G. *Molec. gen. Genet.* **119**, 139–152 (1972).
21. Roberts, J., Roberts, C., Craig, N. & Phizicky, E. *Cold Spring Harb. Symp. quant. Biol.* **43**, 917–920 (1978).
22. Ogawa, T. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **43**, 909–916 (1978).
23. Weinstock, G., McEntee, K. & Lehman, I. *Proc. natn. Acad. Sci. U.S.A.* **76**, 126–130 (1979).
24. Sussman, R., Resnick, J., Calamne, K. & Baluch, J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5817–5821 (1978).
25. Walsh, N. & Stocker, B. A. D. *J. Bact.* **137**, 830–838 (1979).
26. Mount, D. W., Walker, A. C. & Kosel, C. L. *J. Bact.* **116**, 950–956 (1973).

Nitrogen fixation gene (*nifL*) involved in oxygen regulation of nitrogenase synthesis in *K. pneumoniae*

Susan Hill, Christina Kennedy & Eugene Kavanagh

ARC Unit of Nitrogen Fixation, University of Sussex,
Brighton BN1 9RQ, UK

Richard B. Goldberg & Robert Hanau

Department of Biology, Temple University, Philadelphia,
Pennsylvania 19122, USA

The enzyme complex nitrogenase, which reduces N_2 to NH_4^+ , involves two redox proteins, both irreversibly damaged by O_2 (ref. 1). Enzyme activity therefore requires anaerobic conditions, a source of reductant and a large amount of ATP (~16 ATPs per N_2)^{2,3}. In both aerobic and facultative anaerobic N_2 -fixing bacteria, nitrogenase synthesis is regulated by O_2 and NH_4^+ , but in the aerobes there are also processes to protect the enzyme from O_2 damage^{4,5}. The mechanisms of repression by O_2 and NH_4^+ seem to be independent in the organisms so far examined^{6–8}. In the facultative anaerobe, *Klebsiella pneumoniae*, O_2 was shown to repress nitrogenase synthesis in an NH_4^+ -constitutive strain⁸. The fusion of the *Escherichia coli lacZ* gene into each transcriptional unit of the nitrogen fixation (*nif*) gene cluster in *K. pneumoniae* has facilitated studies with O_2 , because expression from the various *nif* promoters results in an O_2 -stable product (β -galactosidase). Notably, the *nifHDK* operon (the nitrogenase structural genes) was more sensitive to O_2 repression than the *nifLA* operon (regulatory genes)⁹. The characterization of mutants, reported here, indicates the involvement of a *nif*-regulatory gene product in the mechanism of O_2 control of nitrogenase synthesis.

The *nif-lac* fusions constructed by Dixon *et al.*⁹ with phage Mud(Aplac), in which *lac* is expressed from *nif* promoters, could not be directly used for the isolation of regulatory mutants, as this phage can re-transpose to new locations at high frequency (about 10^{-4} per generation; R. Dixon, personal communication). Therefore our approach was guided by the following argument. All previously described *nifL* mutations (except *nifL2265*) have been shown to be polar on the *nifA* gene¹⁰, the product of which is required for expression of all other *nif* transcriptional units⁹. *NifL2265* showed a partial trans dominance to pRD1 (ref. 11) (a plasmid carrying the entire wild-type *nif* cluster) which is consistent with a modified *nifL* gene product acting as a negative effector. Also, reversion to *Nif*⁺ by strains carrying Mu insertions in *nifL* probably occurs by the imprecise excision of Mu, implying that *nifL* product is not necessary for nitrogenase synthesis¹⁰. Hence revertants of *nifL*

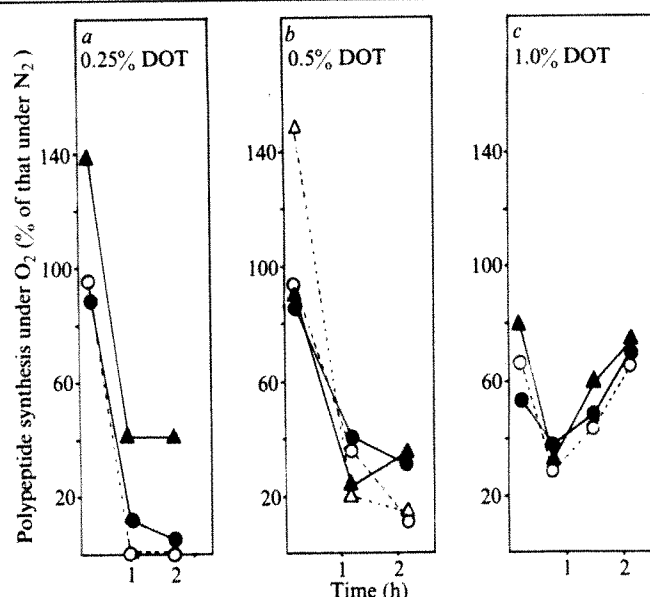


Fig. 1 Influence of DOT on synthesis of *nifH* ($\blacktriangle$), *nifD* ($\bullet$), *nifK* ($\circ$) and *nifJ* ($\triangle$) gene products in the wild-type UNF50231 (a, b) and the mutant CK400 (c) strains. Results show the extent of *nif* gene product synthesis in organisms exposed to O_2 , expressed as a percentage of that occurring in a parallel culture in N_2 . Synthesis of *nif* polypeptides was quantified as the proportion of total polypeptide synthesized, and in cultures in N_2 the initial proportions for each *nif* product ranged as follows: *nifH*, 0.07–0.19; *nifD*, 0.05–0.13; *nifK*, 0.05–0.10; and *nifJ*, 0.02. Total protein synthesis was not markedly affected by exposure to O_2 . Similar experiments were performed on strains UN4357 and UNF753 and *nif* polypeptide synthesis after about 2 h exposure to 0.5% DOT compared with that occurring in N_2 was respectively as follows: *nifH*, 85 and 76%; *nifD*, 96 and 72%; *nifK*, 30 and 36%; and *nifJ*, 180 and 78% (similar results were obtained for strain UN4374); these values are significantly higher than the ones for the wild type shown in b of 35, 33, 12 and 13% for *nifH*, *nifD*, *nifK* and *nifJ* polypeptides respectively. In all experiments cultures were derepressed in an NH_4^+ -free medium²² in N_2 at 29°C. A 20-ml sample was transferred anaerobically to the N_2 -flushed oxystat vessel, equipped with a calibrated O_2 electrode mounted in a side arm and a magnetic stirrer. Immediately after transfer air was slowly introduced into the N_2 supply to the vessel until the desired DOT was reached (never taking longer than 12 min), when the oxystat (Western Biological) was switched on. The DOT value, maintained ($\pm 20\%$) by control of the gas mixture, occasionally drifted, but the time during which the DOT fell below 0.25% (the lowest DOT tested and found to repress nitrogenase synthesis) never exceeded 8% of the total time of an experiment. A control 20-ml sample was stirred in N_2 in a vessel similar to and alongside the oxystat vessel in a water bath maintained at 30°C. Samples (1 ml), removed at time intervals, were labelled with ^{14}C amino acids and protein subsequently analysed and quantified by SDS-polyacrylamide gel electrophoresis and densitometry of autoradiograms as described previously⁸.

mutants might yield strains in which the *nifA* gene product is synthesized but *nifL* gene product is inactive or lacking. Thus we screened *Nif*⁺ revertants of strains carrying mutations in *nifL* for the failure of O_2 to repress synthesis of nitrogenase polypeptides.

Initially we investigated the influence of dissolved O_2 tension (DOT) on nitrogenase synthesis in wild-type strain UNF 50231. As expected, a DOT of 0.5%, which inhibited expression of β -galactosidase in strains carrying a *nifH::lac* fusion, repressed synthesis of nitrogenase polypeptides, as did a DOT of 0.25% (Fig. 1a, b). Besides the products of the *nifH*, *nifD* and *nifK* genes, the rate of synthesis of the *nifJ* gene product also declined (Fig. 1b).

A *Nif*⁺ revertant (strain CK400) derived from a *nifL* point mutant (*nifL2265*) showed some escape from repression when exposed to 0.5% DOT. The kinetics of nitrogenase polypeptide synthesis in 1% DOT. (Fig. 1c) confirmed that O_2 repression

was less severe than in the wild-type strain UNF50231 (Fig. 1a, b). Two other *Nif*⁺ revertants (UN4357 and UN4374; W. J. Brill, unpublished), derived from two strains carrying *Mu* insertions in *nifL*, behaved similarly (see Fig. 1 legend) and synthesis of the *nifJ* product also escaped repression in all three *Nif*⁺ revertants (see Fig. 1 legend).

Both the *Nif*⁺ and O₂ constitutive properties of the revertant CK400 were due to a second mutation (*nif2400*) mapping within the *nif* region; a His⁺*Nif*⁺ transductant (UNF753) of a (*his-nifM*) deletion mutant infected with phage P1 propagated in CK400 showed escape from O₂ repression (see Fig. 1 legend). To facilitate further physiological and genetic experiments, the *his*-linked mutations in CK400 were transduced into both *nifH::lac* and *nifN::lac* fusion strains. As expected, for both fusion alleles β -galactosidase activity could be detected by blue colour formed by transductants growing on NH₄-free X-gal plates incubated in air (see Fig. 2 legend), and derepression of

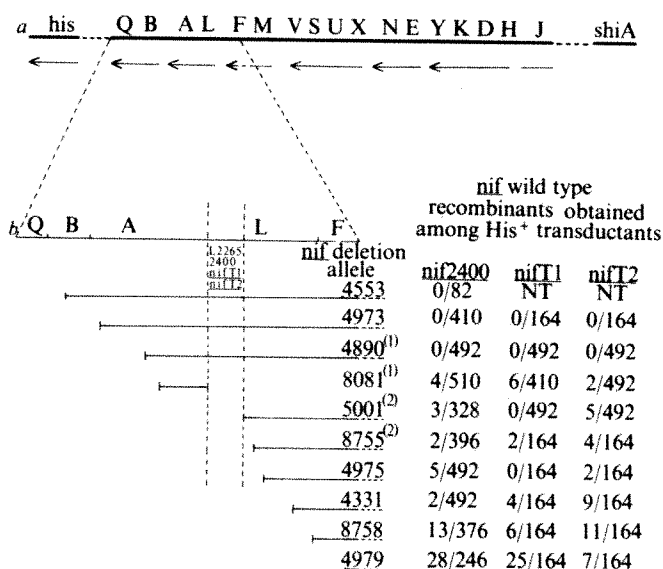


Fig. 2 Genetic mapping of O₂ constitutive *nif* mutations. *a*, Current map of the *nif* genes^{10,23,24}. The lines beneath the map indicate probable transcriptional units and the arrows show the direction of transcription. *b*, The O₂ constitutive *nifH::lac* strains described in the text were made His⁺ and used as recipients in phage P1Cmclr100- or P1Cmclr100-mediated transductions from the set of His⁺ *nif* deletion mutants^{10,24} shown above. His⁺ transductants were selected and scored for *lac* expression on solidified NH₄⁺-free medium²² agar plates containing 30 μ g X-gal (5-bromo-4-chloro- β -D-galactoside) and 50 μ g serine per ml incubated at 29 °C in air; the mutants appeared blue whereas wild-type *nifH::lac* recombinants were colourless on aerobic plates. Possible wild-type transductants were always checked for retention of *nifH::lac* on anaerobic X-gal plates where they turn blue. If wild-type recombinants appeared among transductants, then wild-type DNA corresponding to the position of mutations responsible for O₂ constitutivity was not deleted from the donor deletion strains. If no such transductants appeared, then the mutations lie within the deleted *nif* segment. The number of transductants needed to be scored for *lac* phenotype (~400) to detect any wild-type recombinants was indicated from a similar deletion mapping experiment with a strain carrying the *nifL2265* allele in which 3.5 and 0.5% of His⁺ transductants were *Nif*⁺ using the *nif8755* and *nif8081* deletion alleles, respectively. In the pairs of deletion mutations marked (1) and (2) the relationship between the *his*-proximal ends is uncertain. The *his*-distal end of *nif8081* is near the *nifAL* intergenic region and probably within *nifL* as no *nifA* mutations have been found to map in the deletion interval defined by the ends of the *nif8081* and *nif8755* mutations. The placement of *nifL2265* within this deletion interval conflicts with other reported mapping experiments^{10,24}; however, repeated experiments where His⁺ transductants were scored for *Nif*, rather than being directly selected on NH₄⁺-free medium which is less efficient, consistently gave this result. The *nifT*⁻ mutations T1 and T2 were derived from strains Kp5161-3 and Kp5160-3, respectively^{13,25}. NT, not tested.

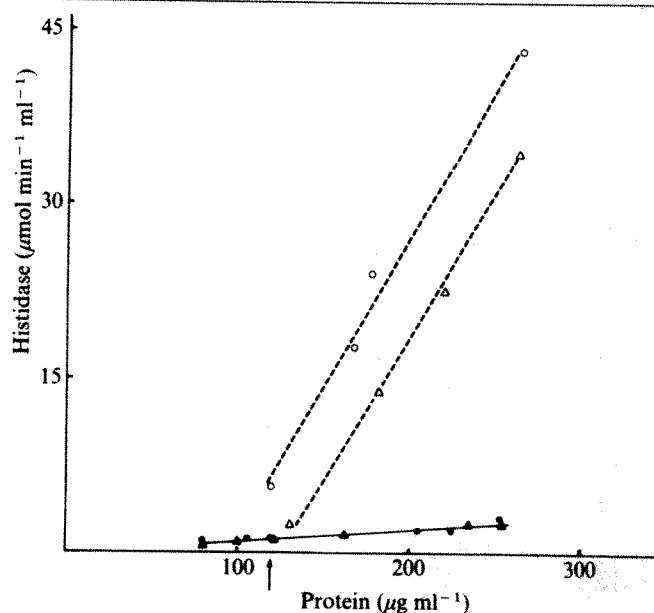


Fig. 3 Influence of an anaerobic to aerobic shift on the expression of histidase in the wild-type strain KG7035 (triangles) and the mutant UNF753, described in the text (circles). Cultures were grown in anaerobic conditions (closed symbols), and when the total protein concentration had reached 120 μ g ml⁻¹, indicated by the arrow, the culture was divided and half was transferred to an aerobic environment (open symbols)¹⁹.

β -galactosidase also occurred in liquid NH₄-limited cultures either maintained in a DOT of 0.5% or bubbled with air (Table 1). In contrast, the parent fusion strains were colourless on aerobic X-gal plates and failed to derepress β -galactosidase in 0.5% DOT or air bubbling (Table 1). The apparent insensitivity of β -galactosidase derepression to O₂ in the *nifH::lac* fusion strain carrying the *his*-linked mutations in CK400 was not due to an altered respiratory activity; the potential respiratory rate¹² and the O₂ demand in a DOT of 0.5% were the same in the mutant and parent strains (data not shown).

Three other suspected *nif* regulatory mutations including a previously mentioned *Nif*⁺ revertant of a *nifL::Mu* strain (UN4357) and two mutants designated '*nifT*⁻'¹³ also allowed synthesis of β -galactosidase in air when linked to *nifH::lac* or to *nifN::lac* (Table 1). Significantly NH₄⁺ repressed *nifH::lac* and *nifN::lac* in all four mutants, and O₂ repression was restored in each *nifH::lac* strain by the introduction of the plasmid pRD1 (Table 1).

The blue phenotype on NH₄⁺-free X-gal indicator plates incubated in air was used to locate the *nif2400* and the two *nifT*⁻ mutations, and deletion mapping experiments (Fig. 2) indicate that all three mutations lie within the *nifLA* operon and probably within the *nifL* gene. The *nifT*⁻ mutants were isolated as *Nif*⁺ colonies from a glutamine synthetase mutant with impaired regulation for *nif* and *hut* (histidine utilization)¹³. Hence the *nifLA* operon seems to be involved in the control of *nif* expression by both O₂ and, as suggested previously⁹, the regulatory proteins of nitrogen assimilation¹⁴.

Several enzyme systems in the Enterobacteriaceae are apparently regulated by O₂ (refs 14–18) and in *K. pneumoniae*, *hut*¹⁹ and *nif*⁸ are controlled by both O₂ and NH₄⁺. We therefore examined the effect of the *his*-linked mutations in CK400 on *hut* expression; in both wild-type and mutant strains, O₂ was needed for histidase derepression (Fig. 3). Furthermore, these mutations did not alter the transient influence of O₂ on glutamine synthetase levels¹⁹ (data not shown).

Thus five different strains had *nif* mutations that relieved O₂ repression of the *nifHDK* and, where tested, the *nifEN* and the *nifJ* transcriptional units. In at least one of the strains (CK400) O₂ regulation of another operon outside *nif* was unaffected. The mutations mapped in the same deletion interval that includes the *his*-proximal end of the *nifL* gene. The *trans* dominance of

Table 1 Influence of O₂ on β -galactosidase derepression in strains carrying *nif-lac* fusions

Strain no.	Relevant chromosomal marker in addition to <i>recA56</i> and <i>lac</i> fusion allele <i>nifH2783</i> or <i>nifN2790</i>	Plasmid	Units of β -galactosidase activity after 6 h in		
			0.5% DOT	N ₂	N ₂ plus NH ₄ ⁺ (19 mM)
With <i>nifH2783</i>					
UNF762	None	None	12 ± 9	120 ± 2	1 ± 1
UNF777	<i>nifL2265</i> , <i>nif2400</i> (mutations from strain CK400)	None	156 ± 13	110 ± 30	5 ± 2
UNF767	<i>hisD</i>	pRD1	16	889	5
UNF798	<i>hisD</i> , <i>nifL2265</i> , <i>nif2400</i>	pRD1	27	864	2
			Air bubbling	N ₂ bubbling	N ₂ bubbling plus NH ₄ ⁺ (23 mM)
UNF762	None	None	9 ± 0	129 ± 17	ND
UNF777	<i>nifL2265</i> , <i>nif2400</i>	None	52 ± 3	71 ± 3	ND
UNF954	<i>hisD</i> (<i>nif</i> mutations from strain UN4357)*	None	335 ± 55	191 ± 17	2
UNF954	<i>hisD</i> (<i>nif</i> mutations from strain UN4357)*	pRD1	15 ± 2	622 ± 36	2
UNF951	<i>hisD nifT1</i> ⁻	None	129 ± 4	231 ± 10	2
UNF951	<i>hisD nifT1</i> ⁻	pRD1	5 ± 0	452 ± 52	2
UNF952	<i>hisD nifT2</i> ⁻	None	100 ± 12	261 ± 8	3
UNF952	<i>hisD nifT2</i> ⁻	pRD1	9 ± 0	366 ± 8	3
With <i>nifN2790</i>					
UNF979	None	None	6 ± 0	51 ± 1	2
UNF982	<i>nifL2265</i> , <i>nif2400</i>	None	90 ± 10	99 ± 7	3
UNF984	<i>hisD</i> (<i>nif</i> mutations from strain UN4357)*	None	68 ± 8	88 ± 12	3
UNF986	<i>hisD nifT1</i> ⁻	None	80 ± 8	87 ± 3	2
UNF988	<i>hisD nifT2</i> ⁻	None	33 ± 3	124 ± 9	3

NH₄⁺-grown cultures were suspended in an NH₄⁺-free medium and β -galactosidase derepression was followed as described previously⁹. Cultures were maintained in 0.5% DOT as described in Fig. 1 legend. ND, not determined.

* Strain UN4357 is a Nif⁺ derivative of strain UN2464 carrying *nifL5037::Mu*¹⁰.

wild-type *nif* and the relative insensitivity to O₂ of *nifL* expression⁹ suggest that the *nifL* product is involved as a negative effector in the O₂-regulatory mechanism specific to nitrogen fixation. Conversion to a negative effector may involve processes associated with O₂ metabolism rather than the direct interaction with the O₂ molecule. The mechanism of action of such a negative effector could be: (1) to prevent *nifA* expression; (2) to inactivate the *nifA* product; (3) to bind at each *nif* operator/promoter; (4) to promote *nif* mRNA breakdown; (5) to prevent *nif* translation; or (6) to promote *nif* gene product breakdown. The first seems unlikely because in strains carrying a *nifA::lac* fusion (where *nifL* is expressed) β -galactosidase derepression was not inhibited in 0.5% DOT (unpublished results). The sixth is improbable as synthesis of a hybrid polypeptide is unlikely in strains carrying the *nifH::lac* or the *nifN::lac* transcriptional fusions^{9,20,21}.

This work was supported in part by NSF grants PCM76-01915 and PCM79-22122 to R.B.G. We thank Dr S. Kumar for help in *nifL2265* mapping experiments, Drs F. Ausubel, C. Elmerich, W. Brill and D. MacNeil for strains, Professor J. R. Postgate and Drs M. Merrick and B. E. Smith for reading the manuscript, and L. Batsleer for technical assistance.

Received 18 August 1980; accepted 21 January 1981.

- Eady, R. R. & Postgate, J. R. *Nature* **249**, 805-810 (1974).
- Ljones, T. *FEBS Lett.* **98**, 1-8 (1979).
- Andersen, K., Shanmugam, K. T. & Valentine, R. C. in *Genetic Engineering for Nitrogen Fixation* (ed. Hollaender, A.) 95-110 (Plenum, New York, 1977).
- Kennedy, C. & Eady, R. R. in *Nitrogen Assimilation of Plants* (eds Hewitt, E. J. & Cutting, C. V.) 73-84 (Academic, London, 1979).
- Robson, R. L. & Postgate, J. R. *A. Rev. Microbiol.* **34**, 183-207 (1980).
- Postgate, J. R. *et al.* in *Biological Metabolism of Inorganic Nitrogen and Sulfur Compounds* (eds Bothe, H. & Trebst, A.) 103-115 (Springer, Heidelberg, in the press).
- Bergersen, F. J., Turner, G. L., Gibson, A. H. & Dudman, W. F. *Biochim. biophys. Acta* **444**, 164-174 (1976).
- Eady, R. R., Issack, R., Kennedy, C., Postgate, J. R. & Ratcliffe, H. J. *gen. Microbiol.* **104**, 277-285 (1978).
- Dixon, R. *et al.* *Nature* **286**, 128-132 (1980).
- MacNeil, T., MacNeil, D., Roberts, G. P., Supiano, M. A. & Brill, W. J. *J. Bact.* **136**, 253-266 (1978).
- Kennedy, C. *Molec. gen. Genet.* **157**, 199-204 (1977).
- Hill, S. J. *gen. Microbiol.* **93**, 335-345 (1976).
- Ausubel, F., Riedel, G., Cannon, F., Peskin, A. & Margolskee, R. in *Genetic Engineering for Nitrogen Fixation* (ed. Hollaender, A.) 111-128 (Plenum, New York, 1977).
- Leonardo, J. M. & Goldberg, R. B. *J. Bact.* **142**, 99-110 (1980).
- Picinoty, F. *Biochim. biophys. Acta* **64**, 111-119 (1962).
- Spencer, M. E. & Guest, J. R. *J. Bact.* **114**, 563-570 (1973).
- Harrison, D. E. *Adv. microb. Physiol.* **14**, 243-314 (1976).
- Fimmel, A. L. & Haddock, B. A. *J. Bact.* **138**, 726-730 (1979).
- Goldberg, R. B. & Hanau, R. J. *Bact.* **141**, 745-750 (1980).

- Cassadaban, M. J. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4530-4533 (1979).
- Cassadaban, M. J., Chou, J. & Cohen, S. N. *J. Bact.* **143**, 971-980 (1980).
- Cannon, F. C., Dixon, R. A., Postgate, J. R. & Primrose, S. B. *J. gen. Microbiol.* **80**, 227-239 (1974).
- Puhler, A. & Klipp, W. in *Biological Metabolism of Inorganic Nitrogen and Sulfur Compounds* (eds Bothe, H. & Trebst, A.) (Springer, Heidelberg, in the press).
- Merrick, M. *et al.* *J. gen. Microbiol.* **117**, 509-520 (1980).
- Ausubel, F. M., Margolskee, R. F. & Maizels, N. in *Recent Developments in Nitrogen Fixation* (eds Newton, W., Postgate, J. R. & Rodriguez-Barrueco, C.) 347-356 (Academic, London, 1977).

Erratum

In the article 'Generation of F₁ hybrid cytotoxic T lymphocytes specific for self H-2' by K. Nakano *et al.*, *Nature* **289**, 559-563 (1981), Table 1 contained some errors. The correct version appears below.

Table 1 Genetics of target determinants for the lytic activity of B6D2F₁ anti-B6 CTLs

Group no.	Strain	Target cells				% Specific lysis (effector:target cells = 40:1)
		H-2 haplotype	K	I	SD	
1	B6	b b b b b b b b				36.0
	B10	b b b b b b b b				32.4
	C3H.B10	b b b b b b b b				34.3
2	A	k k k k d d d d				9.0
	B10.A	k k k k d d d d				8.2
	DBA/2	d d d d d d d d				5.8
3	B10.D2	d d d d d d d d				4.4
	C3H	k k k k k k k k				12.7
	B10.A(4R)	k k b b b b b b				23.2
4	B10.A(2R)	k k k k d d d b				26.0
	HTG	d d d d d d d b				38.0
	B10.A(5R)	b b b k k d d d				8.4
5	B6 × DBA/2	b b b b b b b b				8.6
	B6 × C3H	d d d d d d d d				
	B10 × B10.A(2R)	b b b b b b b b				11.1
6	B10 × B10.A(2R)	k k k k k k k k				
	B10.A × B10.A(2R)	b b b b b b b b				31.1
	B10.A × B10.A(5R)	k k k k d d d d				14.1
7	B10.A × B10.A(5R)	k k k k k d d d				
		b b b k k d d d				13.4
		b b b k k d d d				

MATTERS ARISING

'Lipidic particles' are intermembrane attachment sites

THE nature of lipidic 'particles' and 'pits' in freeze-fracture replicas of liposomes has been much debated. Recently, Miller¹ presented arguments that the particles and pits represent conical deformations of membranes, rather than inverted micelles sandwiched between membrane leaflets—a model repeatedly suggested by Verkleij *et al.*²⁻⁴. We have also observed lipidic particles and pits in liposomes of mixed egg phosphatidylcholine (PC)-soybean phosphatidylethanolamine (PE)⁵ and mixed PC-PE-cardiolipin, and our results essentially agreed with those of others¹⁻⁴. Moreover, we have more convincing evidence for the deformation model as presented below. The evidence has led us independently to arrive at a similar model⁶ to that proposed by Miller.

We have obtained cross-fractured profiles of the particles and pits in multilamellar vesicles. As shown in Fig. 1a, most of the remaining halves of the fractured particles and pit are seen as interlamellar contact points, although a few single particles (open arrows) are also observed. The profile of the lipidic particle is clearly conical, with the base extending to 200 nm. They are considerably larger than the mean size of proteinaceous intramembranous particles, as seen in Fig. 1b. Here, in a PC-PE vesicle containing human erythrocyte membrane protein, large conical lipidic particles are clearly seen at the fracture line as contact sites between adjacent membranes. Proteinaceous particles are hemispherical with diameters of ~10 nm. As the lipidic particles and pit are really cusp-shape distortions, they should be named as such instead. When multilamellar vesicles containing <30% PC in PE were quenched at room temperature, rows of cusps on the fractured faces were frequently observed. However, cusps are not found in large single lamellar vesicles (verified by negative staining) of the same lipid mixture. This provides independent evidence that intermembrane attachment is required for the formation, stabilization and alignment of the cusps. Our results give definite support for the cusp model¹ in preference to the inverted micelle model².

This work was supported by NIH grant GM 28120 and American Cancer Society grant BC-248.

S. W. HUI
THOMAS P. STEWART

Electron Optics Laboratory,
Biophysics Department,
Roswell Park Memorial Institute,
Buffalo, New York 14263, USA

1. Miller, R. G. *Nature* **287**, 166-167 (1980).
2. de Kruijff, B. *et al.* *Biochim. biophys. Acta* **555**, 200-209 (1979).
3. Verkleij, A. J., Mommers, C., Gerritsen, W. J., Leunissen-Bijvelt, L. & Cullis, P. R. *Biochim. biophys. Acta* **555**, 358-361 (1979).
4. Verkleij, A. J., van Echteld, C. J. A., Gerritsen, W. J., Cullis, P. R. & de Kruijff, B. *Biochim. biophys. Acta* **600**, 620-624 (1980).
5. Hui, S. W., Stewart, T. P., Yeagle, P. C. & Albert, A. *Arch. Biochem. Biophys.* (in the press).
6. Hui, S. W., Stewart, T. P., Boni, L. T. & Yeagle, P. L. *Science* (in the press).

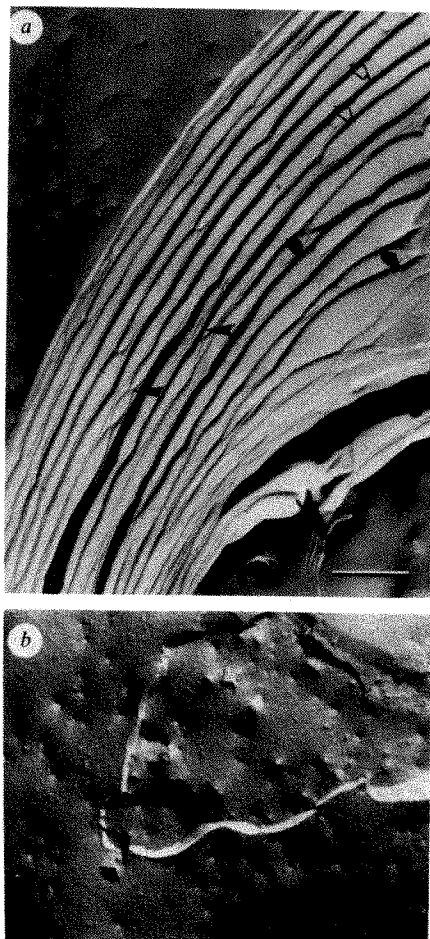


Fig. 1 Freeze-fracture electron micrographs of multilamellar vesicles of 25% PC in PE (a), and with added crude human erythrocyte membrane protein (smaller particles) extracted with Triton X-100 (b). Lipidic particles and pits are seen as intramembrane attachment sites in cross-fractured areas (arrowhead), some of which have developed extensive contacts (solid arrows). Open arrows indicate occasional conical particles on one bilayer alone. Experimental procedures are given elsewhere^{5,6}. Scale bar, 500 nm.

VERKLEIJ AND DE KRUIJFF REPLY—We believe that 'lipidic particles' represent inverted micelles formed either within a single bilayer^{1,2} or at the nexus of intersecting bilayers during fusion of membranes³⁻⁵. We have presented^{2,3} evidence for this interpretation and

against the lipidic particles being local bilayer curvatures as inherent to the intermembrane attachment side (IMAS) model as proposed by Miller⁶ or the similar cusp model proposed by Hui and Stewart in their letter. These arguments and new facts supporting our hypothesis can be summarized as follows:

(1) The crucial point ignored by the IMAS model is that the lipidic particles have been exclusively observed in lipid systems in which at least one of the lipid species prefers the hexagonal phase II (H_{II}). In this latter phase the lipid molecules are cone shaped and oriented with their polar head group towards an aqueous channel of a small radius. Any model of the lipidic particle must take into account this favourable organization of H_{II} -type lipids, a criterion met in the inverted micelle model. This requirement, however, is not fulfilled with the IMAS model. For instance, in the cardiolipin-lecithin system⁶, cardiolipin will be present at the top of the strongly curved bilayer where it can only fit if the molecules have an 'inverted cone' shape. However, cardiolipin assumes a cone shape in the presence of Ca^{2+} .

(2) In line with this argument, a direct morphological association has been observed between the H_{II} phase and the lipidic particles^{4,7}, indicating a close relationship of these structures. Furthermore, we have found in some systems^{3,7} three-dimensional arrays of lipidic particles which probably represent an agglomerate of inverted micelles. An analogous example of structural relationships between micellar and hexagonal phase can be found in another area of lipid polymorphism in that lipids can adopt both micellar and hexagonal H_I phases in water, as, for instance, does lysophosphatidylcholine⁸.

(3) In the IMAS model, the bilayer must be strongly curved as the lipidic particles can be very small (60-130 Å in diameter)². Such a strong bilayer curvature is energetically highly unfavourable because the minimum curvature of the bilayer in the most strongly curved system (sonicated vesicles) is at least 250 Å in diameter.

(4) The IMAS model ignores the extensive ^{31}P NMR studies on the lipidic particle-containing systems^{2,9}. The profound hysteresis of the isotropic peak in the ^{31}P -NMR spectrum of these systems can be best explained by a model in which inverted micelles are located in rows at the fusion site of two bilayers¹⁰.

In conclusion, we feel that these arguments strongly suggest that lipidic particles represent intrabilayer inverted micelles. Also, the observation that in systems such as that described by Hui and Stewart the particles are relatively large and cone shaped can be interpreted as the

presence of inverted micelles at the fusion site of two well separated bilayers¹¹.

A. J. VERKLEIJ
B. DE KRUIJFF

*Institute of Molecular Biology,
State University of Utrecht,
PO Box 80.063,
3508 TB Utrecht,
The Netherlands*

1. Verkleij, A. J., Mommers, C., Leunissen-Bijvelt, J. & Vervegaert, P. H. J. *Th. Nature* **279**, 162-163 (1979).
2. de Kruijff, B. et al. *Biochim. biophys. Acta* **555**, 200-209 (1979).
3. Verkleij, A. J., Mommers, C., Gerritsen, W. J., Leunissen-Bijvelt, J. & Cullis, P. R. *Biochim. biophys. Acta* **555**, 358-362 (1979).
4. Verkleij, A. J., van Echteld, C. J. A., Gerritsen, W. J., Cullis, P. R. & de Kruijff, B. *Biochim. biophys. Acta* **600**, 520-624 (1980).
5. de Kruijff, B., Cullis, P. R. & Verkleij, A. J. *Trends biochem. Sci.* **5**, 79-81 (1980).
6. Miller, R. G. *Nature* **287**, 166-167 (1980).
7. van Venetie, R., Leunissen-Bijvelt, J. & Verkleij, A. J. *Biochim. biophys. Acta* (submitted).
8. Luzzati, V., Reiss-Husson, F., Rivas, E. & Gulik-Krzywicki, T. *Ann. N.Y. Acad. Sci.* **137**, 409-413 (1966).
9. Cullis, P. R. & de Kruijff, B. *Biochim. biophys. Acta* **559**, 399-420 (1979).
10. Cullis, P. R., de Kruijff, B., Hope, M. J., Nayar, R. & Schmidt, S. *Can. J. Biochem.* **58**, 1091-1100 (1980).
11. Verkleij, A. J. et al. (in preparation).

Fossil charcoal and the palaeoatmosphere

IT has been proposed by Cope and Chaloner¹ that oxygen levels in palaeoatmospheres may be assessed by the presence or absence of charcoal in the geological record. Although oxygen is required for burning, it is not necessary for charcoal formation. When wood is subjected to high temperatures, oxygen-bearing gases are evolved, yielding an oxygen-depleted graphitic network (charcoal)². Charcoal formation is favoured by low oxygen concentrations at the solid surface and forms in anoxic conditions at temperatures above 200 °C. In the remote past, however, heat from wood fires, sustained by atmospheric oxygen, was probably the dominant agent for charcoal formation. Thus, the coupling of atmospheric oxygen concentration to charcoal formation may not be as direct as has been proposed.

That the occurrence of charcoal is determined by the oxygen concentration required for carbon monoxide or methane to burn, as Cope and Chaloner presume, has not been demonstrated. Depending on configuration, ignited wood samples may not continue to burn in atmospheres containing as much as 13-21% oxygen³. At lower concentrations the most likely source of heat for charcoal formation would not be available.

It is not known whether palaeoatmospheric pressures were identical to present-day values. The proportion of oxygen required for ignition depends on the pressure⁴, but the dependence has not been investigated for wood. If inert gas is added to air, as in submarines, the partial pressure of oxygen and consequently

respiration are unaffected, but the proportion of oxygen may be reduced sufficiently to prevent a match from burning⁵.

While applauding the cross-disciplinary approach of Cope and Chaloner, we caution that the effects we have described should be considered before this potentially valuable probe of palaeoatmospheric composition be used. With refinement, this might be extended to explore fascinating problems, such as atmospheric implications of expanding Earth models and the effects of oxygen and carbon dioxide partial pressures on photosynthesis⁶⁻⁸ in the past.

F. R. S. CLARK

*Fire Research Section,
National Research Council
of Canada,
Ottawa, Canada K1A 0R6*

D. A. RUSSELL

*Paleobiology Division,
National Museum of Natural
Sciences,
Ottawa, Canada K1A 0M8*

1. Cope, M. J. & Chaloner, W. G. *Nature* **283**, 647-649 (1980).
2. Browne, F. L. *Report 2136* 18-19 (Forest Products Laboratory, Madison, Wisconsin, 1958).
3. Rasbush, D. J. & Langford, B. *Combustion Flame* **12**, 33-40 (1968).
4. Wharton, R. K. *Fire Mater.* **3**, 39-48 (1979).
5. Carhart, H. W. *Nat. Symp. Fire Aspects of Polymeric Materials* (American Chemical Society, June, 1977).
6. Gardescu, J. & Hagar, W. G. *Pl. Physiol.* **65**, Suppl. 6, 12 (1980).
7. Vidaver, W. & Hall, G. *Pl. Physiol.* **65**, Suppl. 6, 13 (1980).
8. Rogers, H. H., Sinclair, T. R. & Heck, W. W. *Pl. Physiol.* **65**, Suppl. 6, 49 (1980).

CHALONER AND COPE REPLY— Although we are aware that charcoal may be produced in a variety of physical conditions, we must stress that the theme of our letter was to examine the palaeoatmospheric implications of our studies of the widespread distribution of charcoal (fusinite) in the geological record. We did not attempt to present a comprehensive survey of the conditions of charcoal formation.

In the natural environment, charcoal may be produced in conditions of both pyrolysis and combustion, but the mechanisms for charcoal formation are few. First, basalt laval flows are known to engulf vegetation and subject the included woody tissues to pyrolysis, transforming these into charcoal¹. Second, in swamp peat accumulations, local temperature rise caused by exothermic decomposition reactions may initiate smouldering which could produce charcoal. Third, and most important, combustion of vegetation by forest fire can result in the production of large volumes of charcoal as a residue. It is the last mechanism which we believe to be responsible for the generation of the greater proportion of charcoalified phytoclasts which occur in ancient sediments. Thus, our suggested implication for past atmospheric composition is not based on the range of physical conditions for charcoal production but on the

restriction (in terms of oxygen level) which can be placed on the occurrence of large-scale natural burning, that is, the combustion associated with forest fire. We placed this restriction at a minimum oxygen level of ~0.3 present atmospheric level (PAL). We acknowledge that it may be possible to produce charcoal in special circumstances with oxygen levels <0.3 PAL. However, at such levels it would not be possible to maintain the combustion mechanism for the production of the large quantity of charcoal preserved in the geological record, estimated to be 4×10^5 – 4×10^7 tons in post-Silurian rocks².

Since the publication of our letter we have learned (K. N. Palmer, personal communication) that the combustion levels of carbon monoxide and methane described in the experiments of Coward and Jones³ were based on pre-mixed flames. In conditions of natural burning the combustion is by diffusion flames, where the minimum level of oxygen required to support combustion is more likely to be ~13% (~0.6 PAL), corresponding to the lower levels of oxygen concentration in the extinction experiments of Rasbush and Langford⁴. Thus, we are able to revise our minimum estimate to ~0.6 PAL, thereby placing further constraints on estimates of oxygen concentration in post-Silurian atmospheres.

The question of increased atmospheric pressures is interesting. Since it is generally accepted that the oxygen in the Earth's atmosphere has accumulated steadily from photosynthetic build-up, any proposed models of palaeoatmospheres should set the partial pressure of oxygen at or about its present level. Presumably, for an oxygen/nitrogen palaeoatmosphere at higher than present-day atmospheric pressure, the implied increased partial pressure of nitrogen would have to be accounted for by increased rates of volcanic outgassing in the past. We do not know of any geological evidence to support or reject this type of model, and therefore can make no further comment on this aspect. In the absence of such data we can only assume that past atmospheric pressures were similar to those of the present day. In such conditions we offer the occurrence of charcoal in the sedimentary record as a valuable geological indicator in any proposed palaeoatmospheric models.

M. J. COPE

W. G. CHALONER

*Department of Botany,
Bedford College,
Regent's Park,
London NW1 4NS, UK*

1. Lockwood, J. P. & Williams, I. S. *Geol. Mag.* **115**, 69-74 (1978).
2. Cope, M. J. in *Organic Maturation of Sedimentary Organic Matters, Preservation Problems and Fossil Fuel Exploration* (ed. Brooks, J.) (Academic, New York, in the press).
3. Coward, H. F. & Jones, G. W. *Bull. U.S. Bur. Mines* **503** (1952).
4. Rasbush, D. J. & Langford, B. *Combustion Flame* **12**, 33-40 (1968).

BOOK REVIEWS

Doing things in the public interest

Eric Ashby

IT TAKES courage to write a book called *Implementing Public Policy*. The author walks in the footsteps of some very distinguished stylists: Bentham and Bagehot in the last century, Wheare and Aron in this century. It was Bagehot who made the point that success in public policy "depends on a due mixture of special and non-special minds — of minds which attend to the means and of minds which attend to the end". Since Bagehot's time the means have become more sophisticated and the end more controversial. The "due mixture" of special and non-special minds sometimes fails to produce the desired result. As a rule the special mind produces a rational case, based on hard data; the non-special mind employs what is called political judgement, which is just a dignified way of saying "hunch". Lord Salisbury, writing over a hundred years ago, made the point so well that there is no room for improvement: "... you should never trust experts. . . . They all require to have their strong wine diluted by a very large admixture of insipid common sense".

Ruth Levitt's book is a variation on this theme. The merit of her book is that she eschews vague generalization about how a decision to do something in the public interest becomes transformed into something actually done in the public interest. Instead she selects a specific example of public policy and uses this as a paradigm for public policy in general. Her specific example is policy for pollution. It is an apt choice: all political parties agree that it is a good thing to abate pollution, so the issue is not bedevilled by party politics. The means for abating pollution are highly technical and cannot be left to politicians or bureaucrats; the "due mixture" has to include scientists, economists and lawyers. And the end, though not entirely free from controversy, is not obscured by ideologies. The choice of pollution, therefore, clears the ground of several difficulties.

Ruth Levitt's technique is to dissect the *Control of Pollution Act, 1974*, and then to examine its implications in four case-studies: toxic waste disposal, the EEC directive on bathing water, noise abatement, and industrial air pollution. In fact — and Ruth Levitt makes this clear — these case-studies involve not only the *Control of Pollution Act* but other laws as well. She describes, with abundant references to recent literature, how the control of dumping of toxic wastes arose in response to sensational reports in the mass media, not in response to the considered recommend-

Implementing Public Policy. By Ruth Levitt. Pp.213. (Croom Helm/Biblio, Totowa, NJ: 1980.) £11.95, \$30.

ations of a committee appointed to examine the problem by the Ministry of Housing and Local Government. The advice of the special mind was disregarded. It was "insipid common sense" on the part of the non-special mind which, in 1972, hustled a Bill through Parliament in record time to control poisonous waste.

Section 17 of the *Control of Pollution Act* is intended to supersede this hasty legislation; but action on this was delayed for over six years, partly because the special minds could not define precisely what toxic waste was. At long last, on 17 November 1980, regulations to control special waste were laid before Parliament. They illustrate one of the themes of Ruth Levitt's book: the difficulty of getting unambiguous and politically viable advice from the special minds. The regulations proscribe the uncontrolled dumping of wastes "dangerous to life". And how do you think "dangerous to life" is defined? Waste is dangerous to life if "a single dose of not more than five cubic centimetres would be likely to cause death . . . if ingested by a child of 20 kilograms' body weight". It is no wonder that there has been a prayer against these regulations in the House of Lords, where the criterion of dangerous-to-life was ridiculed in a spate of not-so-insipid common sense (one noble baroness called it "cock-eyed").

Ruth Levitt's second case-study is another example of the mismatch between the specialist and generalist ingredients of a political decision. The EEC's policy is to harmonize environmental laws in the member nations; so the Commission issues inflexible directives which are intended to apply from the tip of Scotland to the tip of Sicily (and now to the isles of Greece). The directive on bathing water, obviously drawn up with an eye to a crowded resort on the tideless Mediterranean, prescribes mandatory values for some dozen parameters that have to be measured regularly, especially where "the concentration of bathers exceeds a mean value of 10,000 persons per linear kilometer". If the state of British bathing beaches were a hazard to health, such a draconian system of monitoring might be justified; but there is no evidence whatever to support the view that the very great cost of monitoring bathing beaches in Britain according to this

directive is justified. Yet it is now Community Law; an example of policy-making just for the sake of policy-making.

The other two case-studies are equally illuminating. Policy on noise abatement, like policy on the deposition of toxic waste, is hampered by insufficient technical knowledge. Policy on industrial air pollution is not hampered in this way but it is misunderstood. The achievements of the Alkali and Clean Air Inspectorate are in fact one of the great success stories in the control of pollution. But there has never been adequate communication between the Inspectorate and the public. There is provision in the *Control of Pollution Act* (sections 79–83) to empower local authorities to keep registers, open to public inspection, with particulars of industrial emissions. A Royal Commission, a committee set up by the Clean Air Council, and sundry groups of environmentalists have urged that the public have a right to know what is in the air they breathe, but practically nothing has been done; an example of another point Ruth Levitt makes, that there is a gap — sometimes a deliberately unbridged gap — between the passing of a law and its implementation.

The case-studies are valuable and, on the whole, accurate. (Some of Ruth Levitt's statements are not, in my view, correct, but it would be quibbling to dispute them.) What is disappointing about the book is its neglect of the historical dimension of pollution policy, and the omission of any reference to studies on similar public policy done in the USA and Canada. The purpose of the case-studies is to illustrate the general pattern of decision-making for public policy. But this pattern cannot be understood if it is observed without reference to its evolution. Thus the present interpretation of the formula "best practicable means", discussed in Chapter 7 of the book, makes little sense if one does not take account of the early use of the formula in the Towns Improvement Acts in the early 1800s, the criticism of these by de la Beche and Playfair in their pioneer report to the government in 1846, and the new interpretation given to it by the Alkali Inspectorate in the 1870s.

As for reference to work on the other side of the Atlantic, I think Ruth Levitt's concluding chapter on effective public policy would have been greatly improved if she had taken account of the work done by Dorothy Nelkin and her associates in Cornell. Ruth Levitt mentions the role of

public opinion in agenda-setting, but she barely mentions the implications so perceptively analysed in *Technological Decisions and Democracy* (see *Nature* 280, 849; 1979) and other papers by Nelkin.

There is still great confusion about how public policy is made. The present demand for public participation in policy about the environment makes it very important that this confusion should, as far as possible, be cleared away. Ruth Levitt's book is a

serious and useful contribution toward this end. But, as the author herself says in the preface, her idea has not been to make judgements, but rather "to observe, clarify and draw attention to relevant detail". This is what she has done. □

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge. He was a member of the Royal Commission on Environmental Pollution from 1970 to 1973.

Molecular genetics: a language to be learnt

Jiří Novotný

Evolution and Variation of Multigene Families. Lecture Notes in Biomathematics, Vol. 37. By Tomoko Ohta. Pp. 131. (Springer-Verlag: 1980.) Pbk DM 19.50, \$11.50.

GALILEO Galilei said that mathematics is the language in which God wrote the book of nature. Most of science seems to substantiate this idea, but there are exceptions. The whole field of chromosomal DNA study, for example, presents us with a conspicuously non-quantitative picture. Although this feature may contribute to its present-day popularity, it is also rather disturbing. It signals that our contemporary knowledge of gene and chromosome organization (to say nothing of chromosome evolution) is far from a real understanding and is merely superficial at best.

Only partly is this situation due to a lack of quantitative information. Since the advent of DNA cloning and sequencing, volumes of such data are emerging almost daily. What we seem to be lacking, however, is quantitative *understanding*. Quantitative data are no more useful than qualitative if we do not know what they mean and in what context to understand them. The uniqueness of Ohta's book lies in the fact that, for the first time, it outlines a quantitative theory of what happens with the genomic DNA in the course of time.

The book is brief, informal and although it contains some (rather unimportant) errors, it is appealing in its sincere endeavour to provide us with a unified and thorough explanation of diverse phenomena: genetic correlation among and within multigene families governed by unequal crossing-over, amino acid sequence variation of immunoglobulins, amino acid site linkage disequilibrium, gene conversion, accumulation of non-functional genes. This is a difficult goal; in consequence the book is sometimes hard to follow. Compared to the intricacy of the subject itself, it is nevertheless a condensed and understandable presentation which is well worth reading. Ohta's work goes far beyond the pioneering studies of Smith

(*Science* 191, 528-535; 1976) and others in that it deals not only with the evolution of a single chromosome but also with the sexual gene exchange in a population. A key notion of the theory is the identity coefficient: the probability of two genes of the same multigene family being identical by descent. Several different coefficients are defined and compared to analyse different levels of the genome; conclusions drawn from their comparison are indeed interesting. Particularly original and innovative is the analysis of immunoglobulin V regions. The within-species and between-species coefficients are computed separately for the framework and hypervariable sections of the V genes and shown to vary correlatively. Another part of the book gives a non-traditional discussion of a notorious issue, the question of V gene families and subfamilies and their interrelation in the course of evolution.

I have no doubt that anybody who is interested in molecular evolution will find this book stimulating and useful. □

Jiří Novotný is an Assistant Professor in the Cellular and Molecular Research Laboratory of Massachusetts General Hospital and Harvard Medical School.

Great balls of fire?

E. Philip Krider

Ball Lightning and Bead Lightning: Extreme Forms of Atmospheric Electricity. By James Dale Barry. Pp. 298. (Plenum: 1980.) \$29.50, £18.59.

THE remarkable phenomenon of ball lightning has both frightened and fascinated mankind for centuries. A typical ball lightning is a mobile sphere, about the size of an orange or grapefruit, which is produced during a thunderstorm and which remains luminous for one or more seconds. *Ball Lightning and Bead Lightning* is the second book to be published about this subject by Plenum in the past ten years. The first book, *The*

Nature of Ball Lightning by S. Singer, was an attempt to integrate many of the previous papers on ball lightning, both observations and theories, that have appeared over the past several hundred years.

This second work by J. D. Barry reviews the known physical characteristics of ball lightning that have been deduced from visual observations and laboratory experiments, and then attempts to deduce other characteristics and properties which are consistent with these observations. One chapter is also devoted to a discussion of bead lightning, a more familiar phenomenon but one which is still poorly documented in the scientific literature.

The book begins with a brief description of normal lightning and an assumption that ball and bead lightning do indeed exist. Chapter 2 lists all known references to bead lightning, and then presents all published photographs of this phenomenon together with a critical evaluation of their interpretations. The remaining six chapters are devoted to a review of the characteristics of ball lightning, a presentation and critical evaluation of all published photographs that have been reported to be of ball lightning, a review of the laboratory experiments which have attempted to reproduce the natural event and a comprehensive bibliography.

The book is well written and complete at a qualitative level. The principal deficiencies are in the review of normal lightning characteristics and the discussion of bead lightning — the characteristics of normal discharges are quite important if one is to evaluate production mechanisms or if one wishes to compare ball characteristics with a normal channel. No values are given for leader and return stroke currents or their statistics; there is no discussion of continuing currents; and the electric power and energy dissipated by normal flashes is not considered. There is no presentation of time-resolved photographs, which almost always show the streaks of persisting luminosity that may well be a short-duration form of bead lightning. The suggestion that bead effects are most commonly observed during cloud-to-cloud discharges (p.11) is contrary to the cloud-to-ground emphasis in works previously published by lightning researchers. There is also no discussion of whether a simple projection of the three-dimensional channel tortuosity toward or away from the observer could explain the beaded appearance of, say, a long-continuing current.

Despite these shortcomings, the book is a valuable introduction to a fascinating subject and will certainly stimulate research in this area. The collection of over 50 photographs and a bibliography of over 1600 entries are truly exceptional and will serve the scientific community well for years to come. □

E. Philip Krider is a Professor at the Institute of Atmospheric Physics, University of Arizona, Tucson.

Cooperation in limnology, a further achievement of the IBP

C.M. Yonge

The Functioning of Freshwater Ecosystems. Edited by E.D. Le Cren and R.H. Lowe-McConnell. Pp.588. (Cambridge University Press: 1980.) £40, \$95.

THE objective of the International Biological Programme was the furtherance of our understanding of the biological basis of productivity and its relation to human welfare. The study of productivity involved the organization of international studies on land, in the sea and in freshwaters. The last-named posed the greatest problems, both because previous international studies had been slight and because conditions in such waters range so very widely — from oligotrophic highland lochs to tropical alkaline lakes containing dense but specifically restricted algal populations which support equally specialized fish and bird populations.

Because it involved the stimulation of self-financing local enquiries, no section of the programme demanded greater organization. The eventual world-wide participation is indicated in the maps which adorn the end pages of this book which, following earlier publications concerned largely with methodology, represents the final outcome of the freshwater programme or, more concisely, IBP/PF. Under the overall supervision of the Director General, Dr Barton Worthington, especial responsibility rested with the full-time coordinator, Dr Julian Rzoska, to whose dedicated activities, including the particularly close involvement of eastern European countries, this impressive volume is a tribute.

As Professor Livia Tonelli, the last of a number of convenors, writes in the Introduction, "This book is the result of years of a corporate effort by many scientists all over the world to advance the knowledge of production limnology". The truth of this statement is borne out when it is realized that it represents the combined contributions of no less than 43 authors reporting from the 93 widely scattered sites. Not only was interest so widely stimulated but important bilateral projects came into being between such widely separated countries as West Germany and Brazil, France and Chad, Japan and Malaya, and Britain and Uganda.

Freshwaters vary widely in character and consequent productivity; they are also quickly and often profoundly influenced by local human activities with consequences ranging from total toxicity to extreme eutrophication. Yet they also have great potential, realized most fully in Southeast Asia, for the production of fish and prawns for human consumption.

All means of control, alike against damage and towards increased productivity, depend on basic understanding of prevailing conditions. It is this that is the

theme of the book. The chapters range from discussion of the effects of physical and chemical variables to analysis of primary and secondary production. Studies on the significance of organic matter and decomposers are followed by chapters on trophic relationships and efficiencies, then come others dealing with estimations of productivity and consideration of the value of dynamic models of lake productivity. The final

chapter by G.G. Winberg provides a most interesting summary of the results of Soviet freshwater studies.

The overall result with its final 66 pages of references is a volume of unique value to all limnologists and one of the major achievements of the International Biological Programme. □

Sir Maurice Yonge is an Honorary Fellow in Zoology at the University of Edinburgh.

Wooldridge and Linton re-viewed

F. W. Shotton

The Shaping of Southern England. Institute of British Geographers Special Publication, No 11. Edited by David K. C. Jones. Pp.274. (Academic: 1980.) £14.60, \$34.50.

THIS book was commissioned by the Institute of British Geographers to celebrate the fortieth anniversary of the publication of Wooldridge and Linton's *Structure, Surface and Drainage in South-East England* (IBG, 1939). In a preface and ten chapters, eleven authors are involved, some of them more than once. This assembling of the latest opinions of geomorphologists on the evolution of the topography of their own terrain, namely the Thames Basin, the Weald and the Hampshire Basin, will be appreciated by geologists and geographers from a much wider area. The evolution of the form and drainage of this part of England was argued about long before Wooldridge and Linton, as K. Clayton points out. The value of *Structure, Surface and Drainage* lay less in its conclusions than in the way it combined meticulous field observation with scientific interpretation and thus steered geomorphological thought to its present position.

Although it is stated that the book is a memorial to Wooldridge and Linton, only four chapters examine their conclusions critically. Two of these, by D. K. C. Jones and R. J. Small respectively, deal with the development of the landscape during the Cainozoic era, and although Small adds the words "an alternative interpretation" to his title, it is more of an alternative to Wooldridge and Linton than to Jones. Both authors reject two of the main tenets of the original interpretation, which postulated a late Oligocene or early-Miocene "Alpine storm" of earth deformation with subsequent tectonic quiescence while the land surface was reduced to a peneplane, and a marine invasion (the so called Calabrian or late-Neogene transgression) to a height of about 200m. Instead they envisage a

succession of deforming pulses from the early Eocene onwards, with concurrent erosion or deposition.

Two other papers, by C. A. Baker with D. K. C. Jones and by C. P. Green and D. F. M. McGregor, deal with the diversions of the course of the Thames during the Pleistocene. These I found less satisfying, for although the authors differ from several points of the original interpretation, they also reveal how much still remains unresolved. On the age and origin of the Chiltern Drift, surely pre-Cromerian but not inevitably a till, there is no clear consensus.

The other contributions are not so closely related to Wooldridge and Linton but are no less interesting because of that. They deal with the origin of the sarsens, the residual deposits of the North Downs, the palaeohydrology of the Kennet Valley and the effect of periglacial weathering on Chalk. A final chapter by C.P. Green describes the development of geomorphology as it is now understood and practised, and discusses its probable future. Green points out that modern geomorphology, combining as it does the study of land form and drainage, the physical processes by which these are developed, stratigraphy, sedimentology, palaeoenvironment reconstructions and relative and absolute chronology, was taken over by geographers from the geologists. It gives no pleasure to me to read (and half admit) that this was due to "the relative lack of interest in the Quaternary among British geologists". In fact Wooldridge, acknowledged as one of the great pioneers in modern geomorphology, was, by training and early employment, a geologist. Green's chapter emphasizes that geomorphology is now neither simple geography nor geology but a hybrid of both, biology also providing a formative influence. The subject is truly scientific and very much a part of earth science. □

F. W. Shotton is Emeritus Professor of Geology at the University of Birmingham.

CORRESPONDENCE

Continued from page 356

of confusing the origins with the content of a scientific theory⁴. Any student of sociology or philosophy of science is likely to be intrigued by the social framework of the emergence of sociobiological theory over the last decade and its relationship to the New Right ideology⁵. But having explored that link, it is the task of critics of sociobiology to expose what they see as its methodological, *internal* inadequacies (for example, refs 6 and 7).

If sociobiologists want to avoid the charge that they believe that biology is destiny, they should beware of telling magazines that they know "why we do what we do" or entitling their books *The Selfish Gene*, *The Inevitability of Patriarchy* or *On Human Nature*. The trouble is that they want to have their cake and eat it. They imperialize the human sciences (*vide* the first paragraph of *Sociobiology, The New Synthesis*) and are embarrassed by the outcome.

This is why the final chapters of Dawkins' or Wilson's books are so confusing. Having set out the inexorable destiny of genetic predispositions to xenophobia, aggression, patriarchy or whatever, they invoke the possibility of a human conscious prospect of overcoming these predispositions. As Dawkins puts it in his letter, with respect to philandering males "many humans (have) some at least momentary intention of overcoming their polygamous tendencies. Many even succeed in this". For Wilson euphenics may overcome any "hereditary tendency" to acquire xenophobia. Free will, intentions and wishes (or Dawkins' memes) like the US cavalry, come galloping over the horizon in the nick of time to rescue us from our genes.

But *where* does our free will, etcetera come from? How can we be *both* genetically programmed robotic DNA survival machines and have the extraordinary capacity to transcend these programs? The truth is that sociobiological determinism, when challenged, collapses into the weakest sort of Cartesian dualism. For consistent materialists — like Gould, whom Dawkins quotes, and, I hope, myself too — we must argue that wishes, intentions etc. are as much, or as little, "given" by our genes as any other aspect of our human existence. The interesting thing about humans is that the fact that we can change what we do is as much part of our biology as how we do what we do. To give the example used in my review of Wilson — if humans were quadruped and not biped, their social arrangements would be different; that humans are biped is genetically "given", therefore the human social order is genetically given. The point is that such genetic syllogisms are boringly uninteresting about either any of the crucial aspects of differences *between* human societies or the changes that occur *within* any given society. They neither explain nor predict apartheid in South Africa, cultural revolutions in China, Born Again Christianity in the United States, the welfare state or its dismantling under Mrs Thatcher in Britain, still less any of our individual proclivities.

On such questions then, biology has to be silent. It is because self-styled sociobiologists are *not* silent (I challenge anyone to read *The Selfish Gene* and come away *without* a clear impression of Dawkins' views of what biology

has to say about the Welfare State, sexual mores or microeconomics) that their work both trivializes important social and biological questions and is so amenable to neo-Nazi and New Right ideology.

STEVEN ROSE

*The Open University
Milton Keynes, UK*

1. Maynard Smith, J. *Nature* **289**, 742 (1981).
2. Dawkins, R. *Nature* **289**, 528 (1981).
3. Gould, S.J. *Nature* **289**, 742 (1981).
4. Rose, H. & Rose, S. *Commun. Cognition* **13**, 269-283 (1980).
5. Rose, H. & Rose, S. in *Markets and Welfare* (eds Burkitt, B. & Davey, A.) (in the press).
6. Gould, S.J. & Lewontin, R.C. *Proc. R. Soc. B* **205**, 581-598 (1979).
7. Montagu, A. *Sociobiology Examined* (Oxford, 1980).

Conservation sites

SIR — The attack on conservation sites by Muir (*Nature* 12 March, p.82, 173), apart from ignoring that the scientific value is by no means the only value considered by conservationists, is based on several unacceptable assumptions.

In implying that conservation sites are primarily for the preservation of rare species Muir ignores the fact that most have been primarily designated in order to conserve particular vegetation types, associations of organisms, natural or semi-natural ecosystems. The argument that "value to science of any objects or phenomena lies not in themselves but in the information they yield to study. Once this information is recorded and published whatever value remains in the objects or phenomena is of no value to science" is ridiculous on two counts.

First, he specifies no criteria of "value" that are meaningful. He merely transfers the problem of what he means from conservation sites to published information about them. Even if one disregards this problem one is left, on his own arguments, with a strong case in favour of conservation sites. Ecology is still a young science and it has hardly scratched the surface in its study of the vast majority of ecosystem types. The rate of ecological advance does not compare favourably with the rate of loss of certain habitat types. There are, as yet, no vegetation types or ecosystems whose study has been exhausted to the extent that they are of no further interest to ecologists!

I am responsible for the "sites of special scientific interest" that make up the Malham Tarn Nature Reserve, one of the best documented nature reserves in the country. Its scientific documentation started with observations by John Ray in 1671 and has rapidly accelerated since 1947. We now know enough to realize that it will require much more study than that already accomplished before any scientist would be other than foolish to suggest the site was of no further interest to science.

The tragedy of the present rate of loss of sites of special scientific interest is that we know so little of what is being lost. What little we do know strongly suggests that the Nature Conservancy Council have correctly identified the sites of greatest interest to science apart from them being those necessary if we are to achieve the conservationists' aim to convey the maximum diversity of wildlife into the next century. The conservationist is not, as Muir implies, trying to halt evolution. He is

primarily concerned with the perpetuation of a diversity of habitat types. A variety of forest organisms depends on a variety of forest habitats — regardless of the rates of evolution of their organisms.

R.H.L. DISNEY

*Malham Tarn Field Centre,
Settle, North Yorkshire, UK*

DMSO and immunity

SIR — The observations of Pestronk and Drachman¹ that dimethyl sulphoxide (DMSO) reduced anti-receptor antibody titres in experimental myasthenia gravis in rats are of considerable interest. The authors conclude: "It seems likely that DMSO might also be effective in ameliorating other active humoral immune responses. If so, DMSO may provide an effective new treatment for immune diseases mediated by autoantibodies." Lest this statement lead once more to proclamations of DMSO as a wonder drug, several important points need an answer.

The LD₅₀ of DMSO by intraperitoneal (i.p.) injection is between 10 and 11 per kg for rats and mice. Pestronk and Drachman injected 1.0 ml of DMSO into rats daily for 14 days. Assuming a body weight of 250 g for a rat the dose of DMSO for each rat would be approximately 4 g per kg, or 36 per cent of the acute LD₅₀. The injection (i.p.) of this amount of DMSO into rats and mice produces a number of pharmacological and biological effects including an increase in peripheral blood pressure, changes in the oxygen levels, probable anoxia in the spleen and hypothermia^{2,3}. In mice the level of hypothermia is severe, with rectal temperature reaching 33.5°C and recovery after a single injection of DMSO (4.5 g per kg) takes more than 6 h. Hypothermia represents a generalized but reversible toxicity. Many chemicals when injected at levels of approximately 20–40 per cent of the LD₅₀ produce similar effects.

Thus the clearly demonstrated effects of DMSO in decreasing antibody titres illustrated by Pestronk and Drachman may not be specific. In any event the possible therapeutic usefulness of a chemical effective at doses so close to the LD₅₀ is unlikely to prove clinically acceptable.

M. J. ASHWOOD-SMITH,

*University of Victoria,
British Columbia, Canada*

1. Pestronk, A. & Drachman, D. B. *Nature* **288**, 733-734 (1980).
2. Ashwood-Smith, M. J. Thesis, Univ. London (1962).
3. Ashwood-Smith, M. J. *Int. J. Radiat. Biol.* **34**, 101 (1961); *Ann N. Y. Acad. Sci.* **141**, 45–62 (1967).

SIR — The comments of Professor Ashwood-Smith should serve as a useful reminder that the pharmacological actions of dimethyl sulphoxide must be carefully examined. The mechanism of action by which the drug suppresses anti-acetylcholine receptor antibodies in the experimental animal model of myasthenia gravis is not yet understood. Moreover, the reduction in serum antibody levels must be taken into account as a possible undesirable effect if the drug is contemplated for other clinical applications.

ALAN PESTRONK

DANIEL B. DRACHMAN

*School of Medicine,
John Hopkins University,
Baltimore, Maryland, USA.*

9 April 1981

Per Ardua ad Astra (et Terram)

One swallow does not make a summer, but for people yearning to see the winter end, the sight of even a single swallow is bound to seem a cheerful omen. This is the spirit in which the managers of the United States National Aeronautics and Space Administration (NASA) will be glued to their television screens for the next few days, waiting to see whether the new reusable spacecraft, the shuttle, will succeed. With luck, the first version of the shuttle (called Columbia) will be launched on Friday this week, roughly two years later than first planned. Soon afterwards, it should be clear whether the spacecraft has found its way into an orbit about the Earth. Just over two days later, it may have found its way back again, to an airstrip in California. That is the plan. NASA officials will not be the only people biting their nails through this tense period. Several crucial scientific enterprises, the Large Space Telescope in particular, hang on the success of the shuttle. So do several important technical projects, the further development of the international telecommunications network for example. And while it is by no means clear what the Pentagon has in mind for this new means of putting things and people in orbit, the chances are that a success for Columbia this weekend will spark off a round of vigorous jockeying for priority in the launching timetable among the potential users. So much is easily predictable.

It is therefore churlish but also necessary, even at this late stage, to raise the question of what happens if Columbia fails. Ordinarily, if the shuttle were a more conventional spacecraft, the consequences of failure would be negligible. The managers at NASA would bite their lips, utter whatever stoical comments the circumstances of failure required — and set to on the readying of the next spacecraft from the production line. If Columbia fails, they will no doubt follow the same tack but with less conviction. With the budget for the next fiscal year still enmeshed in Congress, it is unlikely that they could escape the thorough investigation of the whole shuttle programme that they have managed to avoid in the past two years. And even if the first flight of Columbia is a success, it is now clear that the development of the shuttle has been too much of a scramble for comfort.

Of the problems that have beset (and delayed) development,

the most serious is that concerned with the device for protecting the spacecraft from heat on reentry into the atmosphere. Recurring troubles with the ceramic tiles intended to accomplish this would have been comic if they were not also almost tragic. It has been forgotten that the tiles are a makeshift substitute for the wonder alloys foreseen when the shuttle was originally conceived of a decade ago, and which have not materialized. The other glaring deficiency in the development programme is the lack of flight experience. A single launching from the back of a Boeing 747 can hardly command confidence. By contrast, recurring problems with the rocket motors, which may nevertheless turn out to be disastrous, are largely a monument to the American passion for optimizing everything — never use a rocket motor off the shelf if there is a chance that a purpose-built motor might be slightly more efficient seems to be the motto. If Columbia fails this weekend, these questions are certain to be reopened.

Others than NASA officials should therefore keep their fingers crossed. The whole device may have been an imperfectly considered technological leap in the dark, but success promises a quantum jump in the utilization of orbits about the Earth. The planned reduction of cost (by roughly a factor of ten for unit mass placed in an orbit) is only a part of the promise. The possibility that the shuttle may be a means of maintaining and repairing satellites already in orbit is potentially more important. Whatever the outcome of this weekend's experiment, there is therefore the strongest possible case for NASA to spend a substantial part of its hard-won budget on a more satisfactory solution than the present of the heat-shielding problem. That cannot be a simple task. Heat-shields based on thermal insulators whose surfaces ablate at high temperatures, widely used to protect the warheads of long-range missiles, are plainly unsuitable for spacecraft intended to be used repeatedly. Solid insulators intended to get rid of heat by radiation are even more susceptible than others to destruction by thermal shock. But mechanical cooling systems based on circulating fluids are probably impractical — their mass would be too great. NASA has already let one contract, seeking an alternative solution. It should quickly cast its net wider. For the shuttle, for all its defects, it is the best American spacecraft in sight.

Who, not Congress, should police fraud?

How often do scientists cheat? How seriously, in the process, is the scientific literature corrupted? And what, if anything, should be done to put a stop to the occasional transgressions that come to light? The United States Congress (see page 435) should not be surprised that it failed to answer these questions in its brief inquiry last week into a few of the cases of wrongdoing that have come to light in the past few months. On the other hand, there is no reason why the research establishment should resent the inquiry or even fear that the outcome will be a novel, intrusive and probably unworkable system of regulation imposed from Washington. There is no reason to suppose that Representative Albert Gore's subcommittee seeks to be a kind of arbiter of ethics in the research laboratory or that it plans to recommend some central system of regulation. The most sinister of the plausible motives for last week's inquiry is that the subcommittee, which plans next to take up more general (and important) issues such as the propriety of arrangements between academic scientists and commercial companies in the field of genetic engineering, the adequacy of agreed procedures for experiments with human subjects and the

like, is seeking to demonstrate at the outset that not all scientists are, like Caesar's wife, above reproach.

The committee's inquiry would have been more productive if it had more clearly thought through the problem of malfeasance in research before dragging its impressive collection of witnesses to Washington last week. It should not have confined its inquiry to biomedical research, and it should have distinguished more clearly between the different kinds of misdemeanours that have from time to time become open scandals. The committee (and those who gave evidence last week) should also have had the wit to recognize that any formal central system for the prevention of misdemeanours is likely also further to inhibit freedom of access to the scientific literature. Journals and their referees are already sufficiently aware of the dangers of having wool pulled over their eyes. More restraint would diminish the chance that heterodox ideas would find their way into the light of day and further tip the scales against innovations.

Plagiarism, the most flagrant of misdemeanours, is paradoxically not as serious a cause for concern as some pretend.

The activities of Dr E.A.K. Alsabti which came to light last year (*Nature* 285, 429 and 286, 437; 1981) were a flamboyant demonstration of how a person can make a mark in the literature with the help of a photocopying machine and a little ingenuity in rewriting the titles of other people's articles. Of necessity, however, success (if that is the word) requires that the plagiarist should be satisfied with publication in relatively obscure journals and that he should pass off as his own only uncontroversial or even dull observations by other people. Journals are not well-placed to spot this kind of transgression — even though some of Alsabti's published papers should have set off alarms in the editorial offices of the journals in which they appeared. But even the most ingenious plagiarist cannot expect a long career. However obscure the journal in which he publishes, and however uninteresting the topic, systematic retrieval systems are now good enough to ensure that those working in the field, including the authors whose work has been plundered, will recognize what has happened within a year or so. In more active fields of research, detection is likely to be even faster. So much is clear from the sensation last year in which one of Professor Philip Felig's collaborators at Yale University was accused of having stolen some text from a paper he had seen as a referee: not unnaturally, the plagiarizing paper was sent to the original authors for review and the whistle was quickly blown. Out-and-out plagiarism, then, is likely to be a nuisance only for librarians and bibliographers. It is exceedingly improbable that working scientists will be misled. And the penalties for the transgressors are severe — they are as a matter of routine expelled from the scientific community.

The falsification of data is necessarily a more troublesome misdemeanour. If a person chooses to send for publication a set of data which is the product of imagination and not experiment, there is very little that journals and their referees can do to identify the deception. Nor would it help if would-be authors were required to send with their manuscripts copies of their laboratory notebooks (even though journals often require that authors should submit evidence to support puzzling conclusions, for example when there are doubts about the statistical analysis of raw data which do not form an integral part of a paper intended for publication). In the ordinary course of business, journals and their referees are not as competent as a reasonably competent investigator to make an assessment of what is involved. The laboratory bench has something in common with the confessional. It must be supposed that a person is telling the truth. If he is not, the penalties are severe, as Dr John Long, dismissed in January 1980 from the Massachusetts General Hospital (see *Nature* 289, 227; 1981) for falsifying published data, has found to his cost.

Unfortunately, cases of falsification rarely emerge as clearly as this. The scientific literature is replete with records of observations which cannot be accommodated within the accepted framework of understanding but which are much more often the result of faulty techniques and mistaken interpretation than of deliberate falsification. And, no doubt, hidden in this mass of data are observations representing genuine challenges to accepted theories. It would be catastrophic if journals habitually declined to give house-room to data that could not be accommodated within the existing body of received wisdom. Although, on such occasions, journals have a duty to enquire closely into the methods by which data have been obtained, in the last resort it is essential that uncomfortable data should be published, not suppressed. If some small part of these data should be fabricated — and it is more likely that false data will be used in the support of orthodoxy, which is just as dangerous — that cannot be the end of the world.

In the nature of things, it is exceptional that a single set of data can overthrow an accepted theory or sustain one that should be discarded. Even dramatic events such as the observation in 1921 that the images of distant stars were deflected at the limb of the Sun as predicted in 1917 by Einstein's theory of gravitation did not still (and have not stilled) arguments about the validity of that monumental work. Similarly, Dr Joseph Weber's entirely well-intentioned claim (in 1968) that gravitational waves had been

detected by equipment based at the University of Maryland and the Argonne National Laboratory, although published in respectable journals (*Physics Review Letters* for example), carried little conviction even though it stimulated several attempts to repeat the observations of a phenomenon that should be observable, although not as strongly as Weber claimed. For what it is worth, the more recent observations by Steele and Gorczynski (see *Nature* 289, 678; 1981) of the inheritance of acquired immunological characteristics in mice deserve a place in the scientific literature even if, in the long run, their interpretation turns out to be different from that now given to them. At least where data bear on general understanding, witchhunting for fabricated data is likely to do more harm than good. Time will tell what data make sense, while the scales are already too powerfully weighted against the heterodox.

So is there nothing for the US Congress to be alarmed about? That, unfortunately, is not the case. Last week's inquiry produced some disturbing evidence of deliberate falsification in the trial and assessment of new drugs. No doubt from the best of motives (a concern that individual patients should have the advantage of the newest drugs), physicians are tempted to include in clinical trials people not eligible under the protocols laid down. Such deceptions will usually lead to the underestimation of the value of a new drug or other treatment, and thus risk the deprivation to future patients of valuable procedures. The safeguard, however, lies in the hands of those who supervise clinical trials — given the understandable temptation, why not demand authenticated medical histories of the patients included in them? Falsification of data bearing on the safety of new drugs, frequently (in the United States) produced by independent contracting laboratories, is potentially a more serious and intractable problem, but here again the remedy is accessible, and administrative. In such circumstances, bodies such as the Food and Drug Administration should (and frequently do) demand access to original laboratory records. Moreover, the agency has sufficiently competent technical staff to carry out a proper assessment of such data. In any case, the laboratories concerned, knowing that their reputations are at risk, have a vested interest in accurate reporting. It should not be beyond the wit of regulatory agencies to ensure that they are not wilfully misled.

So should Representative Albert Gore's committee conclude that there is nothing to be said? That is what Dr Philip Handler was suggesting last week. Unfortunately, that is too simple a view. Although the scientific community is enviably free from wrongdoing, at least as compared with other professional communities (lawyers or politicians, for example), it is too simple to suppose that the peer review system can be relied upon, at least in the long run, to root out all deception. One obvious snag is that acclaimed success, even if based on the over-interpretation of data, tends to breed more success. Grant-making bodies would be less than human if they were able to avoid all such traps. The past few months have also uncovered disturbing evidence that investigators holding substantial research grants are relatively immune from the scientific criticism of their close colleagues, and that they may also be so heavily engaged in the administration of a research programme (including the search for further grants) that they have little time to give to the detailed assessment of the data produced by research programmes in which they are titularly involved. Many of those to whom these descriptions apply will no doubt have mended their ways after learning of the tragic object lessons of the recent past. For the ultimate safeguard of the authenticity of scientific data rests not merely on the peer review system (whose critics necessarily lack first-hand knowledge) but also on full and frank discussion within individual laboratories. The benefits, too often denied, are not merely that a tiny proportion of unauthentic data will struggle with more difficulty into print but that laboratories will in general be made more productive and creative. It would be seemly if the subcommittee in its report urged such steps. It would be ridiculous to go further, and to ask that there should be some formal external process for demonstrating what is already self-evident — that the research profession is as much like Caesar's wife as can be expected.

Universities complain at Pentagon policy

Restrictions on foreign students cause alarm

Washington

Government interference in academic research, ostensibly to control the export of technical data, has stung the presidents of five top universities to protest. The Pentagon is seeking to limit the involvement of foreign research students on certain projects. The universities have also protested, in a letter to Defense Secretary Casper Weinberger and Secretary of State Alexander Haig at attempts to restrict the publication of unclassified research sponsored by the Pentagon.

The university presidents express concern about guidelines issued recently by the department in which defence officials suggest that universities should limit the involvement of foreign students in research related to the development of very high speed integrated circuits (*Nature* 289, 736;1981). They complain that at least one university has been told by the Department of Commerce not to allow certain foreign scholars to participate in its sponsored research activities — a reference to the Export Administration Regulations restricting the transfer of technical data and equipment to communist nations. More action along these lines is expected.

The signatories to the letter are Donald Kennedy, president of Stanford University; Marvin L. Goldberger, president of the California Institute of Technology; Paul E. Gray, president of the Massachusetts Institute of Technology; Frank H. T. Rhodes, president of Cornell University; and David S. Saxon, president of the University of California. According to the text of the letter, not yet officially released by the universities, the full-scale classification of research, where appropriate, is "far preferable" to the application of "restrictive and virtually unenforceable regulations" to universities. For those university activities which remain unclassified, they add, the government should give up all attempts to apply the restrictions until broader issues — which include a possible conflict with the first amendment to the constitution, guaranteeing freedom of speech — have been resolved.

Two sets of export restrictions lie at the heart of the universities' concern. The first is the International Traffic in Arms Regulations, introduced by Congress in the early 1970s to limit the export to any country of technology (including technical knowhow) of potential military value. The

second is the Export Administration Regulations of the Department of Commerce, aimed at limiting the export of technology of commercial value to certain countries, particularly those in Eastern Europe.

Confusion has arisen in universities and scientific organizations — and in the federal agencies themselves — largely because of the difficulty in drawing a clear line between basic knowledge obtained by scientific research and results which, although still described as "basic", have been developed with a particular long-term military or commercial application in mind.

Last year, for example, the regulations were invoked by the Department of Commerce to prevent scientists from Eastern Europe from attending an open meeting on computer bubble memories organized by the American Vacuum Society, one of more than 40 scientific societies which have since applied for exemptions from the export controls for future scientific meetings on the grounds that, since these are open meetings, they should not be considered to involve the transmission of "secret" data.

The concern expressed by the five

university presidents, however, applies more specifically to the conditions that must be observed by scientists in US universities and research institutions, and their freedom to publish the results of their research or to communicate it to non-American scientists through seminars or even informal discussions.

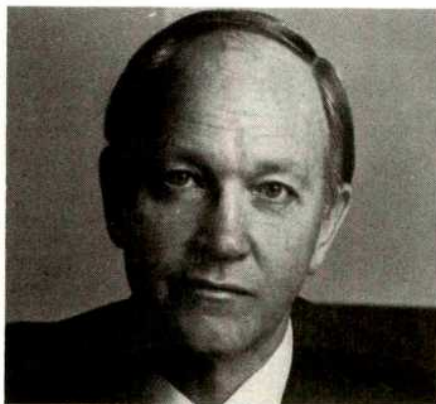
Officials at Cornell University, for example, said last week that in one recent instance the Department of Commerce had placed conditions on a foreign scientist's visit, restricting his access to research at the university, which it had not been prepared to accept, and the visit had been cancelled. In another instance, a research contract proposed by the Department of Defense specified that, although the research was to be unclassified, the department's permission was required before any results were published — a condition which many universities, such as Stanford, say they would find unacceptable.

The five university presidents point out that although the regulations have existed for several years, only recently have they begun to be applied to "traditional university activities". The new interpretation could result in restrictions which

Gowans washes hands of Rothschild

New York

Recent changes in the organization of medical research in the United Kingdom mean that the customer-contractor basis proposed by Lord Rothschild between the Medical Research Council (MRC) and the Department of Health and Social Security (DHSS), "has now been abandoned", Dr James Gowans, director of MRC, told an audience at the Rockefeller University in New York last week.



Gowans — Rothschild off his back

Describing the British approach to medical research funding, Dr Gowans said that after the customer-contractor relationship was introduced in 1972, with substantial funds being transferred from MRC to DHSS, only a handful of new contracts had been awarded. "It became apparent that the customer-contractor

principle was not applicable to medical research, certainly not on the scale envisaged by Rothschild.

Following further discussions, the Rothschild arrangement had been abandoned and the money previously transferred to DHSS had been transferred back to MRC. "We are now back in the situation that we were in before 1972", Dr Gowans said, adding that "to a scientist it is gratifying", since many members of the scientific community had argued strongly against Rothschild's ideas when they first appeared.

MRC and DHSS "both now realize that fundamental solutions to problems such as cancer and coronary heart disease cannot be conjured up by the use of contracts. It is not possible to use bureaucratic devices to accelerate the rate of scientific progress", he added. As part of the new arrangements, MRC had also agreed to undertake more health services research, since the council had been told that much of the work carried out under that heading was very important but not very well done.

Mistakes had arisen in government science policy through a failure to distinguish between areas of research — such as cancer research — where the basic principles were not known and the science could not be substantially accelerated merely by providing additional funding, and areas where the scientific principles were known and additional support would provide more substantial results.

David Dickson

would conflict with the fundamental precepts that define the role and operation of the American university system, they say. Strict interpretation, coupled with the severe criminal penalties that they contain, would be even more restrictive than direct classification of the research.

In response to the letter, whose addressees also include Mr Malcolm Baldrige, the Secretary of Commerce, the three federal agencies involved have apparently told the universities that they are looking closely at the problems raised, but it is expected to be several weeks before a formal reply is agreed upon.

The potential impact of the export control regulations on the academic community is already being discussed by an *ad hoc* committee set up last month by the National Academy of Science's Committee on Science and Public Policy.

The rules are also being closely studied by the National Science Foundation. Dr Donald Langenburg, the Deputy Director, says the foundation is "deeply concerned" about the potential impact of export controls on research, adding that it is "critical" to the freedom of scientific exchange that these are kept to a minimum consistent with foreign policy and national security objectives.

The growing tension between the academic community and the federal agencies over the interpretation of export restrictions is precisely the type of problem for a presidential science adviser, still absent from the White House. **David Dickson**

DNA research guidelines

Dutch get tough

Brussels

The publication by an *ad hoc* committee of revised safety rules for recombinant DNA research has raised some grumbles in Dutch industrial circles. By and large the new guidelines, the conclusions of a year's study of the risks of gene manipulation, follow closely those set by the National Institutes of Health in the United States. But while other European countries are moving towards a relaxation of the rules, the Dutch have tightened up the restrictions on research with pathogenic microorganisms.

Companies such as Unilever and Gest-Brocades, which are carrying out research into the industrial applications of bioengineering are, however, more frustrated about the administrative checks on research. In the Netherlands, DNA research is first vetted by the *ad hoc* committee, although there is no legal obligation to register the research. Local authorities then have the last word over individual experiments and any new research laboratories.

It is up to the committee, the local authority and university advisory groups to interpret the guidelines as they see fit. Industry complains that it is unnecessarily

restrictive to ask permission to carry out each individual experiment. Unilever has taken some of its projects to Belgium and Gest-Brocades has switched some of its resources to its UK laboratories. In the meantime both expect to wait at least a year for approval on new laboratories. The companies feel that local government authorities are much harder on industry than on universities.

The Netherlands is now thought to have the toughest rules in the world. Public distrust of DNA research caused four members of the *ad hoc* committee to abstain from the final vote on the new guidelines on the grounds that there should be a study of the impact of genetic engineering on society as a whole.

The European Community's Economic and Social Committee is taking the same option. Last July the European Commission brought out its own recommendation on the registration of DNA research, which stressed that over the past three years it had become clear that DNA research is not as dangerous as was originally thought. However, before agreement is reached between the Economic and Social Committee and the European Parliament, a public workshop is to be held in May that will bring together representatives from a cross-section of society to discuss the economic, social and ethical implications of such research.

The Commission's recommendation, if it is adopted, would do little more than establish common definitions and provide the basis for a consensus on registration procedures. However, many researchers increasingly feel that arguments about risk assessments are academic. The bulk of research is carried out in the lowest risk class and very rarely in PIII containment conditions. Thus, the slightly stricter rules proposed by the Dutch are expected to have little impact. **Jasper Becker**

Solar Polar mission

Joint project saved?

The European Space Agency (ESA) is "cautiously optimistic" that something will be salvaged from the International Solar Polar Mission, although in a less ambitious form than originally envisaged. The two-spacecraft mission to study the polar regions of the Sun was threatened last month when the US National Aeronautics and Space Administration (NASA) announced that, to meet budget cuts, it would not be building its spacecraft.

Although it is unlikely that NASA will restore the mission to its original status, angry protests and diplomatic pressure by European embassies in Washington seem to have prompted a compromise. According to ESA, that might involve NASA building a cheaper spacecraft — costing \$40 million rather than the \$100 million originally planned.

Under such a scheme, NASA's space-

Congress redispenses

Washington

A congressional subcommittee has now prepared the ground for restoring the United States half of the Solar Polar Mission. Last week, the space science subcommittee of the House of Representatives Science and Technology Committee voted to make cuts elsewhere in the National Aeronautics and Space Administration's (NASA's) budget for the fiscal year 1982; these cuts made it possible for the subcommittee to restore money for projects such as the Solar Polar Mission (which was given an additional \$15 million) without exceeding the total for the NASA budget of \$6,700 million proposed by Mr Reagan. Since the rearranged budget package will not incur additional costs, House Democrats are hopeful that the pro-space science subcommittee of the Senate Commerce Committee, which votes later this month on NASA's authorization, can also be persuaded to accept reinstatement of the solar spacecraft, perhaps on a smaller scale. **David Dickson**

craft, which would lose a de-spun platform carrying a solar coronagraph for direct imaging of the Sun, would be almost identical with ESA's. Two similar spacecraft would allow measurements of spatially and temporally resolved dynamic phenomena, which would not be possible with only one.

ESA's dismay at NASA's decision to abandon its part of the mission was largely due to the fact that it had already spent most of the money for its spacecraft in European industry. NASA, on the other hand, had delayed building its craft after the planned launch date of the mission was put back last year. If NASA is willing to compromise, ESA will agree to a further launch delay to 1986, even though it would increase the costs. The precise details of the compromise will have to wait until Dr James Beggs, NASA's new director, is installed. **Judy Redfearn**

Biotechnology

Canada takes stock

Washington

After six months of intensive study, a task force set up last summer to look at the development of biotechnology in Canada has recommended that the Canadian government should establish a ten-year National Biotechnology Development Plan, with a budget that would rise to about \$50 million a year. Most of this money would go on research and training; exploitation would remain for the private sector, which the task force recommends should receive significant tax incentives.

The task force was set up by Canada's Minister of State for Science and Technology, Mr John Roberts, and was

headed by Dr Maurice Brossard, director of research at the Institut Armand Frappier in Laval, Quebec (*Nature* 285, 431; 1980). Although acknowledging that the extra financial investments which it recommends would be a significant financial burden at a time of fiscal constraint, the task force says that the proposed sum "is a very conservative estimate of required federal expenditure".

Three factors are listed which make this expenditure important — the wide range of potential applications of biotechnology, the importance of biotechnology to Canada's future industrial development and "the almost complete lack of existing Canadian biotechnological capability".

On the last of these, the report is disarmingly frank. The second section of the report, "Biotechnology in Canada: An overview", contains two short paragraphs which refer to a "practically non-existent biotechnological industrial base", a rapidly shrinking federal government research capability and a highly fragmented and unfocused university research effort.

To correct this situation, the task force suggests that the graduate and postgraduate training programmes of the Natural Sciences and Engineering Research Council and the Medical Research Council should be increased by \$3 million and \$1.4 million respectively in the next financial year.

Less predictable is the proposal that Canada's immigration policies be revised "to address the immediate shortage of biotechnological skills which are in high demand worldwide".

The government is expected to endorse the proposal that commercial exploitation be achieved through tax incentives rather than direct government investment. The task force has, however, suggested direct government assistance of up to \$6 million to encourage small and medium-sized biotechnological industries. **David Dickson**

European uranium supply

Australia signs up

Brussels

A large slice of Europe's need for uranium over the next 30 years is to be met by a guaranteed supply from Australian ore. Australia and the European Atomic Energy Community (Euratom) last week initialled a draft of a joint nuclear safeguards agreement in Canberra which should enable Europe to buy uranium from Australia. Negotiations began in 1979 and the Canberra meeting was the fifth round. The agreement has been pursued with great determination by Euratom because Australia, with 20 per cent of the western world's uranium reserves, is very important to Europe's supply strategy.

The basic deal, to run for 30 years, is for Australia to guarantee supply of uranium exports in exchange for European

Community exports of nuclear technology and promises that the radioactive material will not be re-exported or used for military purposes.

Similar agreements are also being negotiated with Canada and the United States. The Canadian deal is a renewal of the 1959 agreement which expired last year and has run into difficulties on the issue of "sensitive" nuclear material (uranium enriched above 20 per cent). Negotiations to renew the agreement with the United States have been held up until the Reagan Administration is firmly in the saddle.

Negotiations with Australia went smoothly until the Australian government decided to impose new conditions. These included a commitment by Euratom member states to ensure an efficient monitoring of reprocessing techniques and to finalize an international plutonium storage programme; a promise not to use or store plutonium extracted from Australian uranium in a way that might lead to proliferation; and that the member states should ask for prior permission on a case by case basis to use the material for research or civil purposes.

In effect, then, the Australians are going further than the Nuclear Non-Proliferation Treaty. The Community negotiating team have been trying to establish precise definitions of the circumstances in which prior permission must be obtained. Does, for instance, using nuclear power to drive submarines constitute a peaceful purpose?

Full details of the new agreement, likely to be similar to the Australian deals with China and Japan, will not be divulged until the final text is agreed. **Jasper Becker**

Fraud in research

Congress asks why

Washington

Dr John C. Long, whose resignation from the Massachusetts General Hospital was forced last year after the discovery that he had fabricated research data, went before a congressional committee last week in sackcloth and ashes. The committee chairman, Representative Albert Gore, complimented him for shouldering full responsibility for his acknowledged "wrongdoing and dishonesty". Dr Long said that he liked his new job (as a surgical pathologist at a Detroit Hospital) — and that he had no plans to return to research.

Dr Long's remarkable statement was one of the few misdemeanours that the government oversight committee of the House of Representatives Committee on Science and Technology had been able to assemble for its two-day inquiry into wrongdoing in science. On the second day of the inquiry, officials from the Food and Drug Administration accused the National Cancer Institute of withholding information about the potential side effects of an anti-cancer drug it was testing. But most of the hearing was taken up with several cases of apparent fraud and abuse

that have recently emerged in the scientific literature and the public press.

The data which Dr Long admitted to fabricating were included in a paper published in the *Journal of the National Cancer Institute* (62, 787; 1979) in response to a critical comment by a referee. But Dr Long also admitted that he had concealed from his colleagues that cell lines purportedly derived from patients with Hodgkins's disease had been contaminated with other cells later shown to have been owl monkey cells (Harris *et al. Nature* 289, 228; 1981).

Nature misrepresented

Washington

In his evidence to the subcommittee, Dr Philip Felig said that two journals, one of them *Nature*, had refused to publish a letter of retraction. Dr Felig said that *Nature's* refusal to publish stemmed from a concern for the personal plight of Dr V. Soman, his collaborator responsible for an act of plagiarism who was afterwards found also to have falsified published data.

After the hearing last week, however, Dr Felig agreed with the editor of *Nature* that his evidence to the subcommittee had been misleading; that *Nature* had agreed to publish the retraction when the audit, then under way, of Dr Soman's laboratory work had been completed; and that at the time (June 1980), he had expressed himself as satisfied with this arrangement.

The chairman of the subcommittee was informed of this conclusion both by the editor and by Dr Felig, and agreed to include a note to this effect in the published record. Dr Felig's intended retraction was in respect of the paper "Regulation of the glucagon receptor by physiological glucagonaemia" by V. Soman and P. Felig, *Nature* 272, 829-832 (1978).

The force of his confession, however, was blunted by the statement by Dr Ronald Lamont-Havers, his former departmental head at the same hospital, that investigations carried out since Dr Long's departure have shown that other data, previously published, may also have been fabricated.

The subcommittee's objective last week was threefold: to estimate the frequency of wrongdoing in biomedical research, to understand the causes of such behaviour, and to decide if anything should be done.

Most witnesses agreed that falsification and plagiarism in the scientific literature is a rare but persistent phenomenon. Dr Donald Fredrickson, director of the National Institutes of Health (NIH), took some pleasure in telling the committee that Gregor Mendel's research results did not stand up to modern statistical analysis.

But even so, the Food and Drug Administration reported that it has disqualified 42

investigators from work with new drugs since 1966, when procedures were devised to deal with violations of the regulations for the administration of new drugs to patients. There have been 20 disqualifications since 1977, when a more formal "Bioresearch Monitoring Program" was set up. At NIH, Dr Raub said that rules covering the grounds on which a research scientist can be barred from receiving further funding had recently been tightened, and that, although no individual had been disbarred in the history of the institutes, at present five instances were on record where a scientist would have questions raised about his or her past conduct if a further research grant application was received.

The committee seemed to accept that known cases of falsification are the tip of only a small iceberg. The suggestion by Dr Philip Handler, president of the National Academy of Sciences, that in the circumstances the committee should turn its attention to more important problems was not, however, well calculated to win friends. The chairman of the committee put Dr Handler firmly in his place with a declaration that falsification entailed the waste of public money and damaged the public reputation of science; and that if the committee was not satisfied that the research community could police itself, it might have to recommend external regulation.

The committee found its position strengthened by testimony from Dr Alexander Capron, executive director of

the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, who recounted the difficulties that the commission was having in obtaining details from the Department of Health and Human Services on the procedures followed in cases of suspected misdemeanour, especially in the treatment of human subjects.

Witnesses differed last week on the motives for falsification. Most of the scientific witnesses held that the pressure of competition for research grants is not an important spur to falsification, but the committee seized on Dr Long's admission that to have shared with his colleagues doubts about the authenticity of his cell cultures would have jeopardized his application for a \$500,000 research grant as evidence to the contrary.

Dr Philip Felig, professor of medicine at Yale University, was closely questioned about the relationship between senior scientists and their colleagues. A case of plagiarism in his own laboratory eventually cost him appointment as dean of the medical school at Columbia University (see *Nature* 285, 429; 1980). He acknowledged that there may be occasions when a collaborator sets out to "please" his senior colleagues, and that the dangers are especially great when the senior person has not himself developed the techniques his collaborators are using. Yale, Felig explained, has now introduced a formal procedure for the independent internal investigation of disputed research results.

Known misdemeanours

Apart from the cases mentioned in the accompanying text, the following is a list of the misdemeanours cited at last week's inquiry by the congressional oversight subcommittee.

University of Kansas (1977). Allegations by two women graduate students that a grant awarded by the National Institutes of Health for anthropological research had been wrongly used for work with human subjects, including the collection of blood samples and genetic counselling. The university established (in 1978) that the work with human subjects had not been cleared with the Institutional Review Board. The investigations (and a civil slander suit by the investigator) continue.

Boston University (1978). Allegations of the falsification of medical records of patients and certain research reports, part of a cooperative study of new cancer treatment financed by the National Institutes of Health (NIH), were substantiated by an internal investigation. The leader of the unit (Dr Marc Straus) resigned on 24 June 1978, followed by his more senior associates, and the Boston University records were expunged from the cumulative records of

the Eastern Collaborative Oncology Group. The investigation continues, but Dr Straus, then at the New York Medical College, Valhalla, was awarded an NIH grant of \$1.3 million over three years in January 1980.

University of California, Los Angeles (1979). Allegations of experimental bone marrow transplants without proper approval from the institutional review board (known as the Human Subjects Protection Committee). The investigations continue.

New York (1978). Inspection by Food and Drug Administration (FDA) revealed "persistent violation" of rules for the administration of novel drugs to human patients by the independent psychiatrist Dr Nathan Kline. Formally debarred by FDA on 13 November 1980.

FDA (various) (1966-81). Total of 42 disqualifications from access to "investigational" drugs, with 18 further cases under investigation. So far, four such investigators have been prosecuted (and one gaoled). Roughly 6 per cent of routine investigations "have led to a more in-depth review". Most violations involve lack of patient consent, inadequate accounting for drugs, violation of protocols and inaccurate (or missing) records.

UK hydraulics research

Deep water ahead

The British government's policy of "privatization" — the sale of public assets — which has already led to the sale of all but a minority stake in the British Aerospace Corporation, now threatens several research laboratories run by the Department of the Environment. The Hydraulics Research Station at Wallingford is the most immediately in danger.

The proposal to make the Hydraulics Research Station self-supporting follows a three-month market survey commissioned by the department and completed last October, in which it was suggested that a potential £1.15 million is available from the private sector. Of this, £0.2 million comes from UK markets, £0.28 million from contracts usually covered by European companies, £0.5 million from offshore projects and the remainder from increasing prices by an arbitrary 10 per cent. The present income is made up of 85 per cent from government and 15 per cent from private contracts.

The market survey suggests that although the UK market is fairly static, the potential for contracts in offshore projects and in the developing countries is considerable. However, a recent survey by the National Maritime Institute foresees a decline in offshore contracts over the next five years. There is also formidable European competition for contracts in developing countries.

The experience of private companies is not encouraging for those at Wallingford. Both the Wimpey Laboratories and the British Hovercraft Corporation have suffered severe financial difficulties. The Water Research Centre, which underwent some sort of privatization following the amalgamation of the Water Pollution Research Laboratory, parts of the Water Resources Board and the Water Research Association, is experiencing similar problems. When it was formed in 1974, the centre was seen by central government as a way of keeping small municipal water undertakings up to the mark. The civil servants involved at first rejected the plan but finally agreed on a package deal. The future of the joint enterprise seemed secure — the centre was financed by its members, with all local and municipal water authorities contributing. In the past year, however, there have been almost 100 redundancies.

Employees at Wallingford now fear that the reputation of the laboratory will be damaged. They also fear that the withdrawal of public funding will inevitably lead to a reduction of background research. However, they see advantages in some sort of halfway "privatization" which would free them from government constraints, particularly on staffing levels, and enable them to take on more contract work. The Government is now

not, however, prepared to consider such a proposal although a meeting has now been arranged for 14 April between the Central trade union and the Department of the Environment. A shadow board is to be set up on behalf of the prospective private company to explore the viability of the enterprise and to prepare employees for privatization — the target being 1 September 1981.

Sara Nash

Europe's directive on drugs

Lords back industry

Proposals from the European Commission to make it easier for entrepreneurs to buy drugs in one country of the European Community and sell them at a profit in another have met with strong opposition. The UK pharmaceutical industry and the House of Lords Select Committee on the European Communities

are at one in condemning the idea as unfair to the companies and unlikely to improve the health services in member states.

The opportunity for this "parallel importing" arises because at the moment price controls for drugs are widely different in the various member states. The European Court of Justice ruled that attempts by big pharmaceutical companies to stop parallel importing were restrictive and in breach of the Treaty of Rome.

Drug manufacturers and importers must obtain marketing authorization from national health services. To make things easier for the parallel importer in the name of free trade, the commission's directive requires the original manufacturer to make available to the importer any necessary documents relating to the specification of the batch of drug being re-sold. This type of freedom of information is anathema to drug companies which may have spent thousands of pounds in obtaining the information in the first place.

In putting the case for the industry, the Association of the British Pharmaceutical Industry (ABPI) argued that moves to bring down the general level of drug prices would not be in the long-term interest of the consumer. And the practice of parallel importing raises problems because importers may change the name and packaging of the drug, with risks that the consumer might not get the correct information about the product.

In the United Kingdom the pricing structure for drugs has kept prices relatively low. In the past few years companies have been allowed to increase prices in line with inflation but for some time "product licences" were given on condition that price increases were kept at about half that rate. As always the industry is keen to stress how expensive it is to develop new drugs and that this should be considered in pricing. So the prospect that the low-price UK market might be opened up to bulk buyers wanting to sell in countries where UK manufacturers are allowed to sell at higher prices is seen as a cause for alarm by both ABPI and the Lords Committee.

The United Kingdom is the most successful exporter in the Community, and last year the British pharmaceutical industry achieved a record £522.9 million surplus of exports over drug imports despite the problems caused by a strong pound. The Lords report sees little reason to jeopardize the well-being of a healthy industry only to give more profits to the middle-men.

The general question of bringing prices into line across the European Community will need more thought than a simple attempt to promote free trade. Both continental and UK manufacturers would favour a uniform system that is subtle enough to take account of both national and international pressures on the companies, and the particular needs and preferences of health services in the member states.

Charles Wenz

Opening up Siberia

Research in steppes

More than 400 scientific research institutes are to be involved in the Soviet Union's new "Siberia" programme which is part of a plan to give this relatively backward area an economic growth rate 30 per cent higher than that of the rest of the country.

Projects will concentrate on the energy sector, according to Academician Valentin Kaptyug, chairman of the Siberian section of the Soviet Academy, with "work of world class" on methods of obtaining synthetic liquid fuel from coal and natural gas. The Baikal-Amur Mainline railway has opened up a new zone for potential development and much of the investment will be in that area. Later more industrialization will take place around the Arctic coast, and in zones around new water resources created by the diversion southward of water from the Siberian rivers.

A food development programme includes proposals from 27 biological and research institutes for the development of new varieties of crops and livestock adapted to Siberian conditions,



Siberia, here we come

"hormonal" methods for increasing crop yields, and pest control methods aimed at leaving minimum toxic residues.

Siberia already produces 50 per cent of the oil and 33 per cent of the natural gas extracted in the Soviet Union. The extraction of fossil fuels in Siberia is expected to go on increasing up to the end of the century, so not surprisingly the new "Siberia" programme also contains work on new prospecting and surveying methods, industrial catalysts and forestry management.

Any practical benefits obtained through the new research programme will be concentrated around "technical production complexes" linking several scores of industrial enterprises. One such complex has already been established, based on the giant power stations on the Angara and Yenisei rivers.

Vera Rich

Tangled policy on drugs

Brussels

The European Court of Justice ruled in 1976 that Hoffman-La Roche were impeding free trade in refusing to give product information to parallel importers. The Italian company Centrafarm had been buying Valium in the United Kingdom and repackaging it for resale under the trivial (not registered) name "diazepam".

The latest proposals from the European Commission are another attempt to limit the monopoly power of large drug companies. Opposition to the Commission comes from most members states, and the European Proprietary Association, representing pharmaceutical companies producing drugs for general sale rather than medical prescription, spoke against the proposals at the Lords committee.

Germany, Italy and the Netherlands are considered high-price countries — on several occasions the European Commission has caught out Italy for supporting unrealistically high prices and court actions have brought down prices in Italy by about 20 per cent.

One effort by the Commission to simplify procedures was the Committee for Proprietary Medicinal Products set up five years ago. The committee can be called upon if a producer wants to put a product on the market in at least six member states, or if there are disparities in the judgements of national licensing authorities. But since it was set up the committee has examined only eight cases.

The adverse comments from the House of Lords seem certain to put another nail in the coffin of the Commission's proposal which has already been languishing for a year in Coreper, the committee of the member states' permanent representatives to the European Community.

Jasper Becker

CORRESPONDENCE

Cancer hazards

SIR — Epstein and Swartz (*Nature* 289, 127; 1981) cited our critique of the report by Bridbord *et al.* on *Estimates of the Fraction of Cancer in the United States Related to Occupational Factors*. Their citation was correct as given, but misleading without our antecedent qualifications. In general, we agree with many others who have reviewed the document that the estimates of occupationally related cancer presented there are probably substantially higher than the data support. In any event, the important question is what reduction in cancer frequency might be achieved by control of occupational exposures. Further epidemiological inquiry should improve those estimates.

REUEL A. STALLONES
TOM D. DOWNS

University of Texas Health
Science Center,
Houston, Texas, USA

Draize test must go

SIR — The news that Revlon is to finance research into developing an *in vitro* alternative to the Draize test for eye irritancy should be welcomed by everyone, not just by anti-vivisectionists (*Nature* 19 March, p.181). The Draize test is notoriously unreliable, yet it has survived basically unchanged for over 36 years. A replacement is long overdue.

As far back as 1971 a survey of intra- and inter-laboratory variation in the interpretation of results indicated that the Draize test was inadequate for assessing eye irritancy even on a simple pass/fail situation. As a result of the study it was recommended that, failing moves to improve reliability, the test had little relevance and should be excluded in regulations as a toxicity test. Despite this, the Draize test is still the only recognized test for eye irritancy.

On scientific grounds alone, the Draize test should have been replaced years ago.

LESLEY SARGEANT

Fund for the Replacement of
Animals in Medical Experiments,
London SW20, UK

IQ put to test

SIR — In his recent *Nature* article¹ N. J. Mackintosh refers to our research demonstrating a relationship between IQ and "inspection time", the latter being a measure assumed to reflect the speed at which a person processes information.

Although we have found inspection time to be correlated with IQ, this relationship becomes statistically significant only for samples with widely varying IQ scores. The correlation coefficient of -0.92 between inspection time and IQ was obtained from a group of 10 individuals whose IQ scores ranged from 47 to 119. We emphasized that this correlation should be interpreted cautiously because of the manner in which the sample was chosen. Furthermore, this relationship was found only with respect to Performance IQ scores, as measured by the Wechsler Adult Intelligence Scale; correlations between Verbal IQ and inspection time estimates were much smaller and not statistically significant². Our study involving 48 subjects, and which found a correlation of -0.80 between Performance IQ

and inspection time³, also included retarded persons. Here again the strong correlation was the consequence of a very wide range in IQ scores. The study involved three groups, each of 16 adolescents; a retarded group where IQ scores ranged from 57 to 81, a group of nonretarded individuals with IQ ranging from 90 to 115, and a third group with scores between 116 and 130. We found no difference in the values of inspection time for the two nonretarded groups (mean = 93 ms, standard deviation = 40 ms and mean = 90 ms, standard deviation = 42 ms). Further analysis of these data (not previously published) confirms only a low correlation between inspection times and IQ for the 32 nonretarded subjects ($r = -0.23$).

At this stage of our research, it is not clear what inspection time is measuring. It is one of a number of measures⁴ developed within a model which attempts to account for capacity limitations, response bias and the caution with which a response is made⁵. The inspection time index suggests that it may prove useful for investigating what limits there are to the speed of human information processing. Insofar as it turns out to be an adequate measure of a sub-process involved in human decision making, it may conceivably be a factor in what might be termed intelligent behaviour, assuming as is widely held that this involves a speed component. To suggest more at this stage, however, seems unwarranted.

T. NETTELBECK
M. LALLY

Department of Psychology,
University of Adelaide (T.N.) and
Research School of Physical Sciences, (M.L.)
The Australian National University,
Canberra, Australia

1. Mackintosh, N. T. *Nature* 289, 529–530 (1981).
2. Lally, M. & Nettelbeck, T. 82, 273–281 (1977).
3. Nettelbeck, T. & Lally, M. *Br. J. Psychol.* 67, 17–22 (1976).
4. Vickers, D. *Decision Processes in Visual Perception* (Academic, London, 1979).
5. Vickers, D., Nettelbeck, T. & Willson, R. J. *Perception* 1, 263–295 (1972).

Why conserve?

SIR — Without seeking to emulate the composer Johann Joseph Fux, who "could not resist being drawn into endless polemics"¹ we feel that some of the points raised by Dr Muir² require critical examination.

The argument that a species or habitat need not be preserved after it has been "competently" investigated raises several questions. First, what criteria can be used to decide whether a particular investigation has, in fact, been competently executed? After the completion of a study by one group of investigators, will it be possible for other workers to confirm or refute their findings, or will the relevant biological phenomenon already have ceased to exist? Second, is it realistic to suppose that everything can be learned about a given phenomenon using presently available experimental techniques? Will future techniques have nothing new to contribute? Third, are biologists concerned only with documenting facts which immediately become historical, rather than remaining current?

Perhaps the most surprising feature of Dr Muir's arguments is that they are utilitarian. Their implication is that a species or habitat should not be permitted to exist for its own

sake, but only if it provides an opportunity for human exploitation. It must be unusual for a professional biologist to express this opinion.

M.F. HEYWORTH
J.S. WAINSCOT

Nuffield Department of Clinical Medicine,
University of Oxford, Oxford, UK

1. Wellesz, E. *Oxford Studies of Composers (1): Fux* (Oxford University Press, 1965).
2. Muir, C. *Nature* 290, 82, 173 (1981).

Lamarck's revival

SIR — The following type of argument has always struck me as unfair to Lamarck and beneath the dignity of a respectable scientist. On the one hand, the emphasis of microevolution and minuteness of changes from one generation to the next, combined with extreme refinement and a plethora of sophisticated tricks (mutation, selection, drift, spatial isolation, genetic isolation, etc.) employed in defence of Darwinism and, on the other, the arrogant vulgarity in discrediting Lamarckism, for example, "why does not a woman whose arm is amputated produce one-armed children?" (ref.1), "can an animal practice being purple?" (ref.2).

The essence of the scientific argument is the validity or otherwise of the Weissman doctrine. It seems logical that resistance to transformation of somatic changes into genetic ones must be high to prevent an organism from the danger of over-reacting to environmental fluctuations and to make it maintain a certain stability. But, in view of the importance of feedback in nature, it seems unlikely that the Weissman barrier is completely leakproof under sustained nonrandom directional stress. Even from the Darwinist viewpoint such complete rigidity and unresponsiveness does not make much sense.

Whether Steel and Gorczynski³ are right or wrong is not so much important in this context. What is extremely important however, is that they are forcing the discussion onto a scientific level which it deserves.

V. KLEMES

National Hydrology Research Institute,
Ottawa, Ontario, Canada

1. *Nature* 289, 631–632 (1981).
2. Savage, R.J.G. *Evolution* 3rd edn, 24 (1977).
3. Gorczynski, R.M. & Steele, E.J. *Nature* 289, 678–681 (1981).

Evolution on trial

SIR — Following the evolution trial in Sacramento, California, thanks are due to scientists who responded to requests from Mr Tyler and myself, acting as his consultant, to be available as witnesses. These were Drs C.R. Berger, B. Dalrymple, R. Dickerson, R. Doolittle, A. Kornberg, J. Kumamoto, W. Laetsch, R. Lemmon, W. Mayer, E. Olson, C. Sagan and D. Wake, evolutionists all, who generously agreed to undergo, if needed, the inevitable frustrations and delays of courtroom trials.

The next "round" will presumably be in Arkansas, where legislature passed a bill, on 17 March, prescribing the teaching of creationism.

THOMAS H. JUKES

Space Sciences Laboratory,
University of California,
Berkeley, California, USA

NEWS AND VIEWS

A tropical Volute shell and the Icarus syndrome

from Jonathan C. Howard

DARWIN tells a little parable in his autobiography¹ which he learned as a very young man, before the *Beagle* voyage, from Adam Sedgwick, the Professor of Geology in Cambridge.

"Whilst examining an old gravel-pit near Shrewsbury, a labourer told me that he had found in it a large worn tropical Volute shell, such as may be seen on the chimney pieces of cottages; and as he would not sell the shell, I was convinced he had really found it in the pit. I told Sedgwick of the fact, and he at once said (no doubt truly) that it must have been thrown away by some one into the pit; but then added, if really embedded there it would be the greatest misfortune to geology as it would overthrow all that we know about the superficial deposits of the Midland Counties. These gravel-beds belong in fact to the glacial period, and in after years I found in them broken Arctic shells. But then I was utterly astonished at Sedgwick not being delighted at so wonderful a fact as a tropical shell being found near the surface in the middle of England."

The point of the parable is that science does not merely consist of making observations but also of ordering them in relation to each other. The tropical shell did not fit into an ordered scheme, and so either the shell or the scheme was wrong. Sedgwick backed the scheme against the shell and won. There would be nothing unscientific about backing the shell, but the quality of support would have to be strong in proportion to the length of the odds against it. No stone could be left unturned to verify every aspect of the find, and eventually a hypothesis would have to be formulated of sufficient generality to accommodate the coexistence of Arctic and tropical shells in the same deposits. Even then the pre-existing scheme would probably not be 'wrong' but merely incomplete. Sedgwick rather overstated the case. New, confirmed and revolutionary discoveries in science very rarely utterly degrade an existing scheme and rebuild its elements in an entirely new pattern. In biology only the theory of evolution by natural selection came close to doing this, and most of the material incorporated into

the new theory was altered only in context rather than form.

The scientific temperament varies, no doubt, along many dimensions. But certainly one of the most interesting is the dimension of credulity. How eccentric does a finding or proposition have to be before it runs off the end of your credulity scale? Are you for the shell or for Sedgwick? Incredulity was my first reaction when I heard by the usual roundabout route that Dr Gorczynski of the Ontario Cancer Institute was reporting the transmission of acquired immunological tolerance through the male. I assumed that the grapevine, with its usual capacity for decreasing the signal-to-noise ratio, had got the sex wrong. Maternal transmission would have fitted comfortably into a well ordered scheme since mother and fetus are known to engage in a rather wide-ranging immunological discourse both before and after parturition; but the grapevine had got the sex right, and here was our tropical Volute shell in the glacial gravel pit. What the grapevine did not say was anything about the genesis of the experiment, and this was a grave omission. Gorczynski's collaboration with Dr E. Steele was conducted with the specific intention of searching for paternal transmission of tolerance, under the influence of Steele's already formulated hypothesis² that RNA viruses might act as vectors for the horizontal transmission of the DNA-coded information from somatic cells to cells of the germ line. According to this hypothesis, transmission through the male is more probable than through the female because male germ cells are formed continuously throughout life.

The phenomenon became public knowledge when the experiments were first published last year³. The level at which I have heard these experiments and their sequel⁴ discussed has varied from gleeful iconoclasm, through critical scepticism tinged with curiosity, to frank disbelief. More upsetting is the charge that the data are suspect because the experiments were

designed and conducted under the influence of a hypothesis of high import. Under these conditions, so the argument runs, the eye is guided more by hope than by the ineluctable evidence of the y-counter. I assumed that this naïve folly had been killed by Sir Peter Medawar's graceful and persistent advocacy of the importance of hypothesis as a guide to useful scientific action, but under the provocation of Gorczynski and Steele's results it seems to have risen from the grave. Since Darwin is often pushed into the ring to confront Steele, I might add that Darwin unequivocally endorsed Steele's logistical procedures when he wrote⁵: "How odd it is that anyone should not see that all observation must be for or against some view if it is to be of any service".

As for the gleeful iconoclasm, it reminds me of the scornful laughter which greeted my Latin master when he muffed an easy declension. He remained an incomparably better classicist than we were. How much of neo-Darwinism are the iconoclasts prepared to throw out on the strength of Gorczynski and Steele's findings? Weissmann's doctrine of the isolation of the germ plasm from somatic influence was a brilliant hypothesis. If it had been badly wrong, genes would not behave in the elegant and predictable way that they do. In the context of Steele's hypothesis, if viral vectors are engaged in the transmission of DNA between cells, and, as implied by Gorczynski and Steele's recent paper in *Nature*⁴, are no respecters of the linkage group, we should have no loci and no alleles. The chromosomes of the germ cells would be a randomized dump of DNA sequences gathered from cells all over the body (see Darwin's *Provisional Hypothesis of Pangenesis* for a nineteenth century expression of this idea)⁶. Every time alleles behave like alleles and linkage groups behave like linkage groups it is another conformation of the generality of mendelian genetics and of the integrity of the germ line, and a further exception to the generality of Steele's hypothesis. However, since neither Weissmann's doctrine nor mendelian genetics predict Gorczynski and Steele's findings, the obligation falls on everyone concerned to find out what

Jonathan C. Howard is at the ARC Institute of Animal Physiology, Cambridge, UK.

kind of exception to the order of things these phenomena are.

It is the mixture of critical scepticism and curiosity alone that is going to resolve this matter. It is a tribute to the vigour of these faculties in the immunological community that this issue of *Nature* carries two manuscripts which go directly to the heart of the scientific problems on hand. First and foremost, are Gorczynski and Steele's results readily reproducible? Brent and colleagues (see p.508) have conducted one of the most serious attempts I have ever seen in immunology to reproduce an experimental protocol in detail. Furthermore, the analysis of the phenomenon is extended in two ways: tolerance was assessed by skin graft rejection as well as the capacity to generate cytotoxic cells *in vitro*, and repeated tests of *in vitro* reactivity were performed on individual animals. There is no escaping the conclusion that, under the conditions of these experiments, the Gorczynski and Steele result has not been confirmed. The explanation for the inconsistency between the two results is now confined to the residual procedural differences between the two series of experiments. Brent and colleagues draw attention to several such differences, only one of which is, in my opinion, important. This difference is in the cell populations used to induce tolerance in the experimental fathers. Gorczynski and Steele's tolerance-inducing protocol was eccentric in the extreme: enormous doses of cells were injected repeatedly into the fathers from birth until the end of the breeding period. Heroically, Brent and his co-workers repeated this anomalous procedure, but they made one mistake. The very first injection of cells, given on the day of birth, consisted of a total of 10^8 composed of equal quantities of bone marrow and spleen cells in Gorczynski and Steele's experiments, and 5×10^7 composed of one part of marrow to ten parts of spleen in Brent and colleagues' experiments. Brent's group is inclined to minimize the significance of this difference, but this seems a foolhardy attitude. Since there was, until Gorczynski and Steele's results were published, no reasonable expectation that any protocol would work, Brent and colleagues cannot dismiss any substantial feature of the successful protocol as insignificant. Curiosity and scepticism are, I am glad to read, carrying these authors still further to the point of reproducing even this aspect of the original protocol in their future experiments.

The importance of this difference in protocol is impossible to assess at present, but there is one crucial point which deserves more emphasis than Brent's group give to it. Amazingly, despite the ruthless régime of tolerance induction, the experimental fathers were apparently not fully tolerant by the measure of their ability to produce *in vitro* cytotoxic activity against the tolerizing parent. In Gorczynski and

Steele's experiments the tolerant fathers were shown to be as tolerant as normal F_1 hybrids. Gorczynski and Steele's *in vitro* induction system was less efficient than that of Brent and colleagues, and weak residual activity may not have been visible, but the difference between apparently fully tolerant and apparently not fully tolerant fathers cannot be considered trivial in an experiment on the transmission of tolerance. Perhaps the difference is real and important, and perhaps it too hinges on the difference in the cellular composition of the first inoculum used to induce tolerance.

When Gorczynski and Steele's first results were published, discussion among colleagues threw up two further kinds of experiment which we felt would explore the generality of the original claim. As self-tolerance is presumably an acquired characteristic, the A homozygous progeny of an $(A \times B)F_1$ father mated to an A mother might show inherited hyporesponsiveness to B alloantigens. Second, since the method of inducing tolerance used by Gorczynski and Steele was so unusual and must have resulted in a high level of chimaerism in the tolerant fathers, we wondered idly what would be the immunological behaviour of the progeny of male $A \leftrightarrow B$ embryo fusion chimaeras with extensive chimaerism in all tissues, ideally including the gonads as well. So rapid and relentless is the march of science that both our lunchtime speculations have already appeared as substantial experimental results. Hasek and colleagues⁷ have done the backcross experiment and McLaren and colleagues (including some of Brent's group) the embryo fusion chimaera experiment (see p.513 in this issue *Nature*). In neither case was any detectable level of specific hyporesponsiveness found in the progeny of the appropriate matings. Again, therefore, Steele's hypothetical mechanism finds itself narrowly constrained by results. If acquired immunological tolerance can be inherited through the male, then the conditions under which this occurs are special and not yet adequately defined. Weissmann and Mendel still provide the general case. We need an explanation for Gorczynski and Steele's tropical Volute shell, but the superficial deposits of the Midland counties are still of glacial origin.

There is much good in this fascinating transaction. A novel and imaginative hypothesis has provoked a series of experiments which would probably never otherwise have been done, certainly not with so much care. The results are conflicting, but there is a *prima facie* case from Gorczynski and Steele's two successful studies for continuing to treat the matter as being of the greatest importance. Even if it turns out, surprisingly, that both sets of positive results have some completely trivial explanation, our perspective will still have been enduringly enriched by the reminder that sex may not be the only way of trans-

ferring DNA between cells.

There is also something bad about what has happened over the past few months, which concerns the fragile and ill articulated matters of tact and style in the presentation of scientific material. Steele's book, *Somatic Selection and Adaptive Evolution*, was unlikely to have turned many heads without some truly original experimental support. It conflates an unacceptably wide variety of experimental findings into a single scheme, and fails in its obligation to apply an intensive critical appraisal in conventional terms to each set of findings in turn. The issue with which it deals, the mechanism of adaptive evolution, is not an entirely open question. Recent evidence for great complexity in the organization of genetic material shows that the beguiling simplicities of the double helix and the genetic code are probably not sufficient to explain all the properties of evolving systems; but this does not mean that all that has gone before can be discarded. Natural selection remains in the pre-eminent position that Darwin and Wallace created for it — as a mechanism for adaptive evolution. One of the most important subsidiary conclusions of genetical research in the last half-century is the recognition that developmental plasticity or individual adaptability to environmental heterogeneity is an evolved character under genetic control. C.H. Waddington⁸ pointed out as long ago as 1942 that genetic modifications to adaptable systems can stimulate the inheritance of acquired characters. In relegating discussion of this fundamentally important concept to a dismissive footnote, Steele does both himself and modern genetics a disservice. It is the current fashion to level popular attacks at the theory of evolution. Twice in the past year Darwin could be found staring gloomily from one corner of the *Sunday Times Magazine*, whilst first Lamark and then Professor Stephen Gould of Harvard University confronted him from the opposite corner. Whether Professor Gould would wish to represent the opposition to classical Darwinism seems doubtful, but the earlier confrontation with Lamark was designed to represent the challenge raised by Steele's work. If Steele is feeling the heat of an intensive critical scrutiny, both of his ideas and of his experiments, it is because, like Icarus, he has flown close to the Sun. □

1. Darwin, C. *Life and Letters* Vol.1 (ed. Darwin, F.) 56 (John Murray, London; 1888).
2. Steele, E.J. *Somatic Selection and Adaptive Evolution* (Williams-Wallace International and Croom Helm, Toronto and London; 1979).
3. Gorczynski, R.M. & Steele, E.J. *Proc. natn. Acad. Sci. U. S. A.* 77, 2871-2875 (1980).
4. Gorczynski, R.M. & Steele, E.J. *Nature* 289, 678-681 (1981).
5. Darwin, C. *More Letters* (eds Darwin, F. & Seward, A.C.) 176 (John Murray, London; 1903).
6. Darwin, C. *Variations in Animals and Plants under Domestication* Vol.2 (John Murray, London; 1868).
7. Hasek, M., Holan, V. & Chuina, J. in *Cellular and Molecular Mechanisms of Immunological Tolerance* (eds Hrabá, T. & Hasek, M.) (Marcel Dekker, New York, in the press).
8. Waddington, C.H. *Nature* 150, 563-565 (1942).

Small is beautiful — portrait of a mitochondrial genome

from Piet Borst and Leslie A. Grivell

PUBLICATION of the complete sequence of the 16,569-base pair human mitochondrial DNA (mtDNA) on p.457 of this issue of *Nature* is a landmark in biology. It comes only 15 years after mtDNA from vertebrates was found to consist of rather small DNA circles¹⁻³, and only 6 years after Sanger and his co-workers developed the sequencing methodology to tackle such large molecules. Together with the virtuoso analysis of mitochondrial transcripts by Attardi's group, reported on pages 465 and 470 of this issue of *Nature*, the sequence provides not only a wealth of information on mitochondrial biogenesis, but also insights into the genetic code and decoding, and into what it takes to make a ribosome, to read messenger RNA (mRNA) and on the minimal requirements for transfer RNA (tRNA) function.

The organization of information in human mtDNA shows an economy never before encountered in nature. Not only does the sequence pack the genes for 2 ribosomal RNAs (rRNAs), 22 tRNAs and 13 assorted proteins so tightly that there are few or no non-coding bases between them, but some of them have been stripped of features that have come to be regarded as *de rigueur* for a respectable gene. In particular, the mRNAs lack non-translated

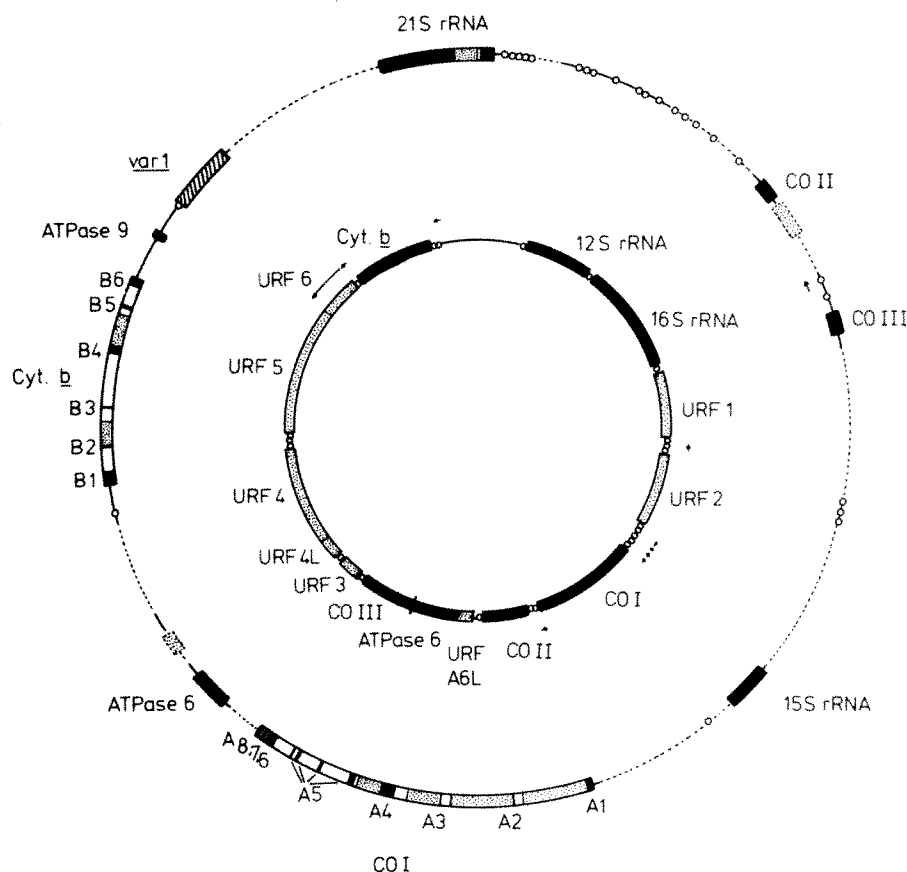
leader and trailing regions. Furthermore, over half of them do not even have a complete stop codon; this only materializes on polyadenylation of the 3' terminus, whereby the terminal U or UA of the gene is converted to the stop codon UAA. Even the transcription-translation machinery used for the expression of human mitochondrial genes has not escaped the drive for economy; features familiar in the genomes of all organisms from viruses to man have been thrown overboard. Thus, transcriptional regulation is a primitive affair, with only one promoter for RNA synthesis on each strand. Consequently, the genome is completely and symmetrically transcribed. Differential gene expression is apparently achieved by a combination of transcription-attenuation, control of the cleavages required to generate transcripts of individual genes and stability of the final transcripts. The rRNAs have been reduced to a size (1,559 and 954 nucleotides) not found outside mitochondria; a 5S rRNA is absent and only 22 tRNAs, of highly unorthodox structure, are used to read all codons. This remarkable feat is accomplished by having some tRNAs read four codons and in these, U is in the first position of the anticodon, permitting U:N wobble.

An important spin-off of the sequence analysis is the finding that the genetic code in human mitochondria differs from that used in other genetic systems: UGA is read as tryptophan rather than 'stop', AGA and AGG as 'stop' rather than arginine, AUA as methionine rather than isoleucine, and AUA or AUU is sometimes used as an initiation codon instead of AUG. The genetic code is therefore not universal.

For the mitochondriacs who spend their days studying mitochondrial biogenesis, the human mtDNA sequence also has an interesting bonus. In addition to the genes for subunits of major membrane proteins like the cytochrome *c* oxidase, apocytochrome *b* and the ATPase complex, which were already known to reside in mtDNA, sequence analysis has unearthed eight URFs (unassigned reading frames, coding for as yet unidentified proteins). There is very little doubt that these sequences represent protein genes because they are conserved in the bovine mtDNA and, moreover, it is highly improbable that such a tightly packed genome would contain non-functional genes. The function of the

Piet Borst and Leslie A. Grivell are in the Section for Medical Enzymology and Molecular Biology of the Laboratory of Biochemistry, University of Amsterdam, The Netherlands.

Fig. 1 Organization of yeast and human mitochondrial genomes. The figure summarizes our present knowledge about the organization of genes on human mtDNA and the yeast *Saccharomyces cerevisiae* strain KL14-4A (~78,000 base pairs). For clarity and to accommodate both genomes within the same format, the human genome has been drawn approximately 2 times 'larger-than-life' relative to yeast. On both DNAs, sequences encoding known gene products are shown as black bars, URFs are indicated by stippled areas and open circles represent a tRNA gene. Transcription is in a clockwise direction, unless otherwise indicated. The cross-hatched region in the yeast genome shows the location of the *var1* genetic determinant, which is involved in the specification of a mito-ribosomal-associated protein. The exact localizations of the DNA sequences encoding this protein are as yet unknown. For yeast, the figure is based on DNA sequence data from several laboratories¹⁸⁻²³, supplemented with unpublished data from our own laboratory (L.A.M. Hensgens, personal communication). Sequenced areas are denoted by a continuous line, unsequenced areas by a broken line. The inventory of genes is thus a provisional one. In the split genes for apocytochrome *b* and COI introns not containing an URF are shown as open bars; exon nomenclature is from refs 19 and 23. The URF in the intron between B2 and B3 in the cytochrome *b* gene has been identified as the C-terminal half of an RNA processing enzyme or maturase (see refs 19 and 24). COI, II, III are subunits of cytochrome *c* oxidase, *cyt. b* is apocytochrome *b*, ATPase 6 and 9 are subunits 6 and 9 of the ATPase complex. URFs are unassigned reading frames. The arrows indicate genes transcribed from the strand opposite the one that encodes most genes.



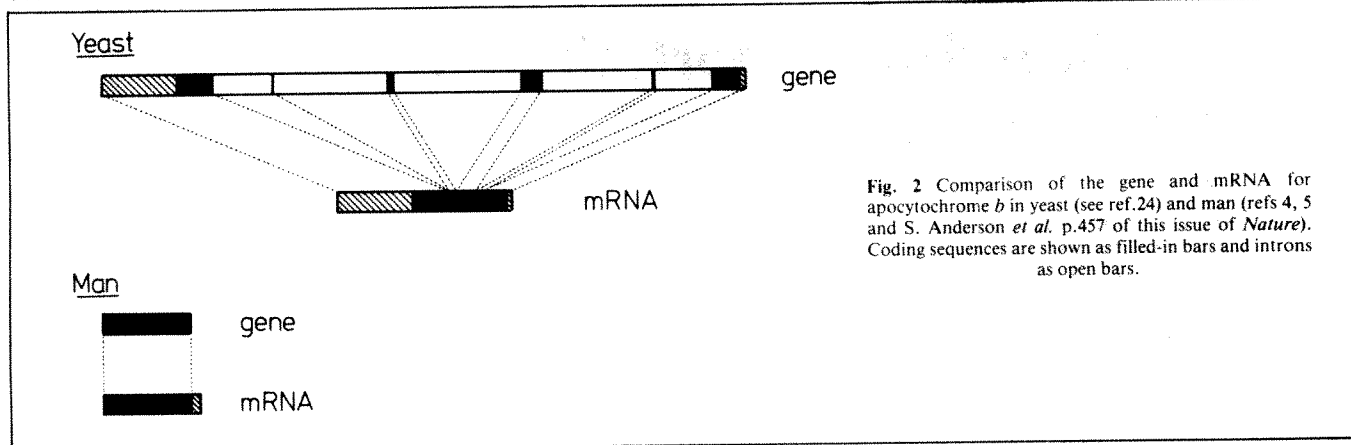


Fig. 2 Comparison of the gene and mRNA for apocytochrome *b* in yeast (see ref.24) and man (refs 4, 5 and S. Anderson *et al.* p.457 of this issue of *Nature*). Coding sequences are shown as filled-in bars and introns as open bars.

proteins corresponding to the URFs remains a matter of speculation (RNA processing? Ribosomal proteins?).

How representative of mitochondrial genetic systems is the human's? The size of mtDNAs from sea urchins, flatworms, insects and man is similar⁴ and the relative positions of tRNAs and other transcripts on human and toad mtDNA is virtually the same⁵. But an entirely different situation is found in yeast, which contains a fivefold larger mtDNA. Figure 1 compares the overall gene order in the two mtDNAs. Yeast contains a series of URFs showing no significant homology with their human counterparts, but has a set of genes for major membrane proteins which are about 50 per cent homologous to their human mitochondrial equivalents. The order and the space in which these genes are arranged in the DNA are, however, completely different. The striking difference in compactness is exemplified by the comparison of the apocytochrome *b* genes, presented in Fig. 2. The coding sequences are almost identical in size, but the yeast mRNA has a leader sequence of 1,000 nucleotides, whereas the human mRNA has none. The yeast gene is split by five introns, the human gene (like all other human mitochondrial genes) is not. The yeast gene has a non-translated 3'-trailing region of about 50 nucleotides; the human gene has no such region and even lacks a complete stop codon. The human mRNA is polyadenylated; the yeast mRNA seems not to be. Obviously there is no pressure for maximal economy on yeast mtDNA. Further evidence of that comes from the long non-coding stretches between genes.

Although RNA processing is prominent in both human and yeast mitochondria, in yeast transcription appears to start from at least five separate promoters⁶ and tRNA punctuation, so elegantly used by human mitochondria in RNA processing (see p.470-475 of this issue), does not have a role. With less pressure for economy one would expect the yeast rRNAs and tRNAs to be less 'petite' and unusual than their human counterparts. This is the case; the rRNAs in yeast mitochondria are about 3,200 (ref.7) and 1,660 (ref.9) nucleotides long and the tRNAs, although lacking some of the common features of tRNAs used in

non-mitochondrial genetic systems⁸⁻¹¹, are less bizarre than the human ones. As in the human, the yeast mitochondria have some tRNAs able to read four codons but they contain more iso-accepting tRNAs, the total gene number now being at least 25, against 22 in human mtDNA. Finally, yeast, like human, mitochondria read the UGA codon as tryptophan rather than 'stop', but they do not share with human mitochondria the unusual assignments of the AUA, AGA and AGG codons. Instead they read the CTN codon as leucine rather than threonine¹², so there is not even a common code in different mitochondria.

Why does human and animal mtDNA look like a 1981 university department, with everything reduced to minimal size without anyone actually being fired, whereas an equivalent set of genes is spread luxuriously over 78,000 base pairs in yeast? We think that the key to an understanding of these differences lies in the rapid evolution of the mtDNA genes that code for rRNAs and tRNAs. For instance, the large rRNA from mouse¹³ and rat (C. Saccone, personal communication) mitochondria share only 68 per cent homology. (Compare this with the 74 per cent homology between the 16S rRNAs of *Zea mays* chloroplasts and *Escherichia coli*¹⁴.) And, at most, human and beef tRNA^{thr} are only 74 per cent homologous. Furthermore, the T ψ C loop, which has a constant seven-base length outside mitochondria, is three bases long in human and eight in bovine tRNA^{thr} (ref.15).

In our opinion this high evolutionary divergence of rRNA and tRNA genes in the mitochondria of closely related species is most easily explained by the low demands made on the mitochondrial translation system. This system has only to produce a very limited number of proteins and usually at a constant relative rate. It may therefore evolve more rapidly than a system which has to produce a large number of proteins at maximal efficiency and subject to various translational controls. Different selective pressures must have pushed animal mtDNA into extreme economy but yeast into maximizing the A+T content. The nature of these pressures is unknown.

If the mitochondrial genetic system can

take more change than other systems, it is possible that further searches in the backstreets of biology may uncover additional surprises. The small mitochondrial genomes of, for example, *Chlamydomonas*, with 15,000 base pairs¹⁶, or *Trypanosoma brucei*, with 20,000 base pairs¹⁷, exhibit different or more extreme solutions to the pressure for economy than human mitochondria. Indeed, the trypanosomes appear to have even smaller rRNAs in their mitochondria (1,230 and 640 nucleotides) than man¹⁷. The study of such extremes may deepen our insight into the minimal requirements for a functional genetic system and provide clues about the evolutionary pathways that might have been followed to develop the more sophisticated mechanisms of gene expression used by present-day bacteria and eukaryotes. □

1. Borst, P. & Rutenberg, G.J.C.M. *Biochim. biophys. Acta* **114**, 645-647 (1966).
2. Sinclair, J.H. & Stevens, B.J. *Proc. natn. Acad. Sci. U.S.A.* **56**, 508-515 (1966).
3. Nass, M.M.K. *Proc. natn. Acad. Sci. U.S.A.* **56**, 1215 (1966).
4. Borst, P. & Flavell, R.A. in *Handbook of Biochemistry and Molecular Biology* Vol.2, 3rd edn (ed. Fasman, G.D.) 363-374 (CRC Press, Cleveland, Ohio, 1976).
5. Rasti, E. & Dawid, I.B. *Cell* **18**, 501-510 (1979).
6. Levens, D. *et al.* in *The Organization and Expression of the Mitochondrial Genome* (eds Kroon, A.M. & Saccone, C.) 265-276 (North-Holland, Amsterdam, 1980).
7. Box, J.L. *et al.* *Cell* **20**, 207-214 (1980).
8. Sor, F. & Fukuhara, H. *C. r. Acad. Sci., Paris* **291**, 933-936 (1980).
9. Berlan, R.E., Bonitz, S.G., Coruzzi, G., Nobrega, M. & Tzagoloff, A. *Nucleic Acids Res.* **8**, 5017-5030 (1980).
10. Newman, D., Pham, H.D., Underbrink-Lyon, K. & Martin, N.C. *Nucleic Acids Res.* **8**, 5007-5016 (1980).
11. Canaday, J., Dirheimer, G. & Martin, R.P. *Nucleic Acids Res.* **8**, 1445-1457 (1980).
12. Bonitz, S.G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 3167-3170 (1980).
13. Van Etten, R.A., Walberg, M.W. & Clayton, D.A. *Cell* **22**, 157-170 (1980).
14. Schwarz, Zs. & Kössel, H. *Nature* **283**, 739-742 (1980).
15. Barrell, B.G. *et al.* *Mosbach Colloq.* **31**, 11-25 (1981).
16. Grant, D. & Chiang, K.-S. *Plasmid* **4**, 82-96 (1980).
17. Borst, P. *et al.* in *The Organization and Expression of the Mitochondrial Genome* (eds Kroon, A.M. & Saccone, C.) 7-20 (North-Holland, Amsterdam, 1980).
18. Dujon, B. *Cell* **20**, 185-197 (1980).
19. Lazowska, J., Jacq, C. & Slonimski, P.P. *Cell* **22**, 333-348 (1980).
20. Tzagoloff, A., Nobrega, M., Akai, A. & Macino, G. *Curr. Genet.* **2**, 149-157 (1980).
21. Bonitz, S. & Tzagoloff, A. *J. biol. Chem.* **255**, 9075-9081 (1980).
22. Nobrega, F.G. & Tzagoloff, A. *J. biol. Chem.* **255**, 9828-9837 (1980).
23. Bonitz, S.G., Coruzzi, G., Thalenfeld, B.E. & Tzagoloff, A. *J. biol. Chem.* **255**, 11927-11941 (1980).
24. Borst, P. & Grivell, L.A. *Nature* **289**, 439-440 (1981).

Thrombosis: a molecular approach to therapy

from P.J. Gaffney

VENOUS THROMBOSIS is thought to be initiated by the thrombin-mediated conversion of circulating fibrinogen to fibrin and the deposition of long fibrin polymers on the vessel wall, with the simultaneous aggregation of platelets and entrapment of other cellular components of blood. There is evidence that fibrin deposition is also involved in the progressive growth of atheromatous lesions, which characterize coronary artery disease. While present-day pharmacological approaches to the alleviation of occlusive vascular disease frequently fail to solubilize the target fibrin without hazardous systemic effects, in this issue of *Nature* (p.505) Smith and his colleagues describe a novel chemical tactic using chemically modified enzymes and activators, which allows the delivery of a fibrinolytic action direct to the target fibrin, leaving the circulating components of the haemostatic system unaffected. Current views about the molecular reactions involved in fibrin formation and destruction allow a better appreciation of the ingenuity of this new approach to thrombolytic therapy.

The generation of thrombin is a central event in the initiation of a thrombus in that this enzyme (1) causes platelets to aggregate, (2) activates fibrinogen to fibrin, and (3) converts an inactive plasma transglutaminase (Factor XIII) to its active form. The latter two reactions are involved in the establishment of an organized insoluble fibrin network which 'cements' the various components of a thrombus together. The interaction of thrombin with fibrinogen involves the removal of small peptides from the dimeric fibrinogen molecule¹, exposing a previously masked dimeric polymerization site on the fibrin monomer² which can attract two fibrinogen molecules in a polymerizing reaction. These latter fibrinogen molecules are converted to fibrin by the action of thrombin, and further accumulation of fibrinogen feeds the multi-branched growth of the fibrin polymer network. Activated Factor XIII catalyses the formation of stable glutamyl-lysyl cross-links³ between the chains of the soluble forming fibrin and this cross-linking process is completed in the insoluble fibrin found in most established thrombi⁴. Thus the *in vivo* formation of fibrin clots is a dynamic process with all the events in a series of reactions being interdependent.

Of necessity there are a number of control mechanisms on clot formation. While inhibitors of thrombin (particularly ATIII) are crucial in the control of thrombin generation, it is generally agreed that the fibrinolytic system protects against fibrin deposition and thus against most

forms of thrombosis. The fibrinolytic system consists of a circulating proenzyme (plasminogen) and a plasma activator (plasminogen activator), the latter being stored in and released on demand from the endothelial cells lining the vessel wall. The interaction of plasminogen activator with plasminogen produces the active serine protease, plasmin, which is very rapidly neutralized by natural inhibitors in the plasma, notably α_2 -antiplasmin⁵. The role of the fibrinolytic system in controlling fibrin generation is best exemplified during disseminated intravascular coagulation, which is a serious complication in a variety of seemingly unrelated conditions (such as cancer, pregnancy and septicaemia). During this form of intravascular coagulation the systemic generation of thrombin is accompanied by the activation of the fibrinolytic system which digests the forming fibrin before it can act as a focus of local thrombus formation on the vessel wall. It seems that as fibrin forms, it orchestrates its own destruction; some of the mechanisms of this orchestration process are now understood (for review see ref.6).

Plasminogen activator and plasminogen bind strongly to forming fibrin while the major fibrinolytic inhibitor, α_2 -antiplasmin, has no such affinity. Thus, this inhibitor, which has such a powerful and immediate neutralizing effect on free plasmin, can only mildly affect plasmin attached to forming fibrin. The preferential binding of plasminogen to fibrin is via a lysine-binding site on the A chain of the plasminogen molecule, which is distinct from that region of the molecule where the catalytic site resides. This lysine-binding site and the active site of plasmin both combine with α_2 -antiplasmin during the rapid neutralization of plasmin. Thus, the protective effect of fibrinolysis against thrombosis seems to be regulated by the localization of plasminogen activator-plasminogen complex on, and the exclusion of inhibitors from, the surface of the forming fibrin⁷.

It is reasonable to propose that the mechanisms described above for systemic intravascular coagulation also apply during the localized coagulation which precedes the occurrence of thrombosis. Should the fibrinolytic system be unable to control such coagulation episodes, it probably cannot digest the consequent organized thrombi containing resistant cross-linked fibrin. There is, indeed, a developing consensus that enzymes from entrapped cells are responsible for the

spontaneous lysis of thrombi⁸. This non-plasmin form of fibrinolysis probably proceeds quite slowly, whereas urgent and effective thrombolytic therapy is required during life-threatening thrombosis, which involves imminent tissue death in vital organs.

The major aim of treatment must be the destruction of the target fibrin in the thrombus, while leaving unaffected the fluid components of the blood. Though two potent fibrinolytic agents, streptokinase (SK) and urokinase (UK), have been in clinical use for the past 20 years, wholly satisfactory therapy has not yet been developed. Urokinase directly converts plasminogen to plasmin while SK forms an equimolar complex with plasminogen or plasmin (SK-plasminogen or SK-plasmin) and this complex, like UK, converts plasminogen to plasmin⁹. While the systemic plasmin so formed is rapidly neutralized by α_2 -antiplasmin, any plasmin generated in excess of the neutralizing effect of α_2 -antiplasmin arguably attacks not only the occlusive thrombus, but also circulating fibrinogen and other components of the haemostatic process. Since these effects are systemic and the thrombus may be localized at one point in the circulation, it is not surprising that both the destruction of fibrinogen and the consumption of α_2 -antiplasmin occur. Damaged circulating fibrinogen and its effect on platelet aggregation have been proposed as major causes of the haemorrhages sometimes occurring during infusion of SK and UK for the relief of occlusive vascular disease¹⁰.

Smith and his colleagues, in their novel approach to the therapy of venous thrombosis, exploited the differential affinities of plasmin for fibrin and α_2 -antiplasmin and demonstrated that acylation of the active site serine allowed plasmin to bind exclusively to fibrin without being inhibited by α_2 -antiplasmin. Likewise, an equimolar SK-plasmin activator complex was acylated and, when infused systemically in experimental animal models, it bound to thrombi without activating the circulating plasminogen. The authors demonstrated the *in vitro* deacylation of plasmin and activator while bound to fibrin and their results in animals indicated a similar deacylation process on the surface of experimental thrombi, causing effective lysis. During these experiments little or no systemic side effects were observed.

This approach provides an attractive alternative to the somewhat hazardous systemic approaches used to date in thrombolytic practice, which are quite often accompanied by the destruction of some essential haemostatic components of

P.J. Gaffney is at the National Institute for Biological Standards and Control, London.

blood with accompanying danger of haemorrhage. The successful targeting of the fibrinolytic action to the fibrin in experimentally induced clots in rabbits and dogs, without overt systemic effects, opens up a new approach to thrombolytic therapy. However, before undue optimism is generated, there is a need to demonstrate these different thrombolytic and systemic effects using acylated enzymes and activators in a whole human system (that is, in a system where a human thrombus is perfused with human blood). The efficacy of this new approach must also be tested in other models of thrombosis which are more

closely associated with the vessel wall. Should these approaches be successful, then a new dimension in the treatment of thrombotic disease may soon be available for testing in clinical trials. □

1. Bettelheim & Bailey *Biochim. biophys. Acta* **9**, 578 (1952).
2. Kudryk *et al. J. biol. Chem.* **249**, 3322 (1974).
3. Pisano *et al. Science* **160**, 892 (1968).
4. Gaffney *et al. Cardiovasc. Res.* **10**, 421 (1976).
5. Wiman & Collen *Eur. J. Biochem.* **78**, 19 (1977).
6. Collen *Thromb. Haemostasis* **43**, 77 (1980).
7. Wiman & Collen *Nature* **272**, 549 (1978).
8. Moroz & Gilmore *Blood* **48**, 531 (1976).
9. Castellino *Trends biochem. Sci.* **4**, 1 (1979).
10. Gaffney *et al. Br. J. Haemat.* **26**, 285 (1974).

Stable clocks and the theory of relativity

from G.J. Edwards

EINSTEIN'S theory of relativity is to my mind the most elegant of all scientific theories. Based on the simple hypothesis that all local, freely falling, non-rotating laboratories are fully equivalent for the performance of all physical experiments (the principle of equivalence), the theory has vast ramifications which extend throughout physics and form a basis for our understanding of the structure of the Universe. The theory has, of course, been something of a theoretician's paradise and an experimentalist's hell, but the past twenty years have seen two developments which somewhat redress this imbalance. These are the ability to hurl sensitive equipment into space and the great improvements of oscillators (clocks) such as masers, lasers and atomic clocks, some of which now have stabilities of 1 part in 10^{14} or so over significant periods of time. The kind of accuracy possible using these aids is exemplified by a recent experiment by Vessot and colleagues (*Phys. Rev. Lett.* **45**, 26, 2081; 1980) which is the most precise confirmation to date of the change in the rate of a clock moving in a gravitational field. The experiment also points the way for future projects which will probe accurately the actual structure of space and time.

In the above experiment (carried out by the Smithsonian Astrophysical Observatory and the George C. Marshall Space Flight Center, US) a very stable clock was launched by rocket to a height of 10,000 km in a near-vertical trajectory. The clock was a hydrogen maser operating at 1,400 MHz and with a stability of 1 in 10^{14} over 100 seconds. After the 18-g launch acceleration, the spacecraft containing the maser went into free fall, allowing about 100 minutes of experimental time, during which the frequency shift in the maser

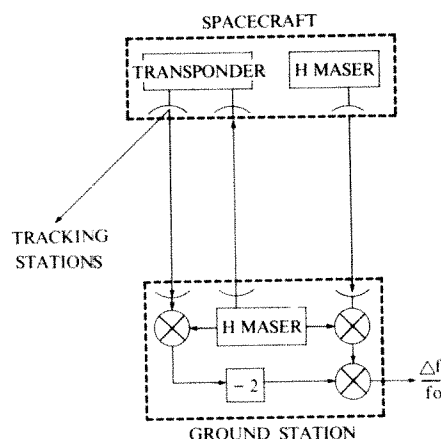


Fig. 1: Simplified schematic diagram of the experiment by Vessot *et al.*

signal radiated to Earth was measured and, after correction, compared with the value predicted by relativity theory (Fig. 1). The main correction was in fact large, being caused by the non-relativistic Doppler shift in radiation from a moving source. Its magnitude was obtained using a separate two-way transponder system between Earth and the spacecraft, as the returning signal contains essentially only the Doppler shift. Finally, the position and velocity of the spacecraft need to be known accurately, and these parameters were deduced from the Doppler-shifted signals received at five tracking stations.

The residual fractional shift, $\Delta f/f_0$, in the maser signal is, according to the theory of relativity, given to first order by:

$$\frac{\Delta f}{f_0} = \frac{1}{c^2} [\phi_s - \phi_e - \frac{1}{2} |V_e - V_s|^2 - r \cdot a]$$

The first term is the gravitational redshift (ϕ_e and ϕ_s are the Newtonian gravitational potentials on Earth and at the maser), while the second term involves the relative Earth-spacecraft velocity and is the famous time-dilation effect for a moving

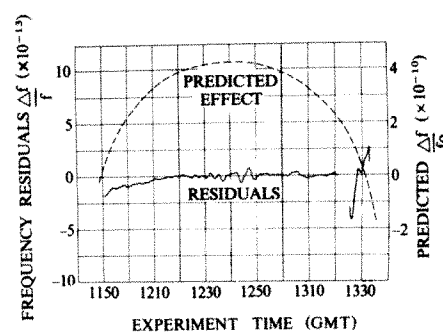


Fig. 2: Frequency residuals and predicted effect during mission.

clock of special relativity. The last term is a residual Doppler shift due to Earth acceleration.

The result of this difficult experiment is indicated in Fig. 2. Clearly the experiment was a great success and the authors quote the overall uncertainty as given by:-

$$\frac{\Delta f}{f_0} (\text{experiment}) = \frac{\Delta f}{f_0} (\text{theory}) [1 + (2.5 \pm 70) \cdot 10^{-6}]$$

The previous best direct verification of the velocity-dependent term is from a measurement of the lifetime of fast-moving muons by Bailey *et al.* (*Nature* **218**, 301; 1977) with an accuracy of 1 in 10^3 . The gravitational shift had previously been confirmed at the 1 per cent level by Pound *et al.* (*Phys. Rev.* **140**, 3B, B788; 1965) using the Mossbauer effect.

Clearly the new result is a significant boost for the theory of relativity and the same is true of other, recent experiments using very stable oscillators. In one of these, Brillet and Hall (*Phys. Rev. Lett.* **42**, 9, 549; 1979) rotated a helium-neon laser working at a wavelength of $3.39 \mu\text{m}$ and locked to a cavity, to check for small anisotropies in space, due for instance to the Earth's motion with respect to the Sun or the cosmic background radiation. They could find no anisotropy at the 1 in 10^{15} level, a 4,000-fold improvement over the best previous attempt. Another experiment by Hafele and Keating (*Science* **177**, 166; 1972) involved flying four very stable caesium clocks around the world in aeroplanes. This experiment is conceptually identical to that of Vessot *et al.*, but less precise. Again, theory and experiment were in agreement.

The experiments outlined above are important indications of the results which will almost certainly be obtained by the use of very stable clocks within the next few years. Besides the general theory of relativity, there are a number of other theories of gravitation which predict different values for the space-time 'curvature' in gravitational fields. The theories differ in their predictions by very small terms which are second order compared with the components of the equation above. Some of these terms have already been measured to modest accuracy, for example by observations on the advance in

the perihelion of the orbit of Mercury and the bending of light rays round the Sun. (For a general review of relativistic, gravitational experiments see Rudenko *Sov. Phys. Usp.* **21**, 11, 893; 1978). However, tests using stable clocks should be able to advance significantly our knowledge in this area.

One such ambitious venture is the solar probe (European Space Agency) which will fly, amongst other things, clocks very close to the Sun, where the gravitational field is very much greater than on Earth. Besides obtaining valuable information on the Sun's structure, the experiment should enable us to distinguish accurately between several of the competing theories. Another experiment proposed by Will (*Metrologia*

13, 95; 1977) is the comparison of two different types of very stable clocks on Earth. Certain (metric) theories of gravitation predict that all clocks, irrespective of their construction, have identical gravitational redshifts while another class of (non-metric) theories predicts small shifts between different types of clock. Such an experiment would require clock stabilities in excess of 1 part in 10^{15} .

All of these tests impose extremely strict requirements on the stability of clocks and great care in the design of the experiments. The reward might be that we discover tiny but significant effects which would imply a change in our idea of gravity as significant as the difference between the Newtonian and Einsteinian theories. □

this process operating in close binary star systems in our galaxy. But at the centre of quasars things are likely to be less simple, the infall of material will probably be more chaotic for example. Before the radiation generated deep down at the centres of quasars is detected by our instruments it will have interacted, that is been scattering absorbed and re-emitted, by material within the quasar. These effects modify the observed continuum spectrum so a simple power law fit that might arise from synchrotron radiation does not usually provide good agreement with the observations. The multifrequency observations now possible using satellites and ground based telescopes have proved very useful, and the various lumps and bumps by which their spectra deviate from a smooth distribution are now the subject of much research.

Another problem which has attracted much interest since the launch of the Einstein X-ray satellite, is the contribution made by quasars to the X-ray background. The first X-ray results lead to the embarrassing conclusion that if the X-ray emission from all such objects were summed out to large distances it would more than account for the observed diffuse X-ray background in the range 1–3 keV. Of the many assumptions which went into this calculation at least one now appears to have been unjustified. The original sample was biased towards radio loud quasars and the contribution of all quasars was based on the assumption that the majority, which are radio quiet, had similar X-ray luminosities. A larger sample of X-ray quasars shows that this is not so. Radio loud quasars tend to be stronger X-ray sources than the radio silent majority. We now believe that quasars contribute somewhere around 40 per cent of soft X-ray background.

A consensus seems to have been reached on a basic model of the line spectra of quasars, both emission and absorption, although it should be borne in mind that the following numbers are order of magnitude at best. In a typical quasar the emission lines arise in a region about 1 parsec in diameter and this volume contains about 100 solar masses of ionized gas. The distribution of the gas is far from uniform and consists of a large number of dense clouds (perhaps 10^{12} of density 10^9 particles per c.c.). These clouds may not be fully ionized and could contain an unknown amount of neutral gas. The cloud model provides a simple approximation and no information is available on the shapes and distribution around the nucleus. It seems that the dominant source of ionization is from the central source and not from collisions between the clouds, although this process is known to produce the emission lines in some nearby active galaxies. One important and recently recognized feature of the ionization

Quasars: what and where?

from Martin Ward

RECENTLY the Royal Astronomical Society was host to a discussion meeting* on Quasars which, although short, did provide an insight into the astronomical communities' current views on the nature of quasars. The talks appeared to divide naturally into two categories — concerning what quasars are and concerning where they are.

The latter question relies heavily on statistical arguments, often a weak point amongst astronomers. A study of pairs of quasars with very similar redshifts shows that, assuming the redshift is a result of the expansion of the Universe, their projected separations in space are between 5 and 30 million parsecs. As clusters of galaxies are thought to be components of 'superclusters' with dimensions of comparable range the observations may suggest that quasars are in superclusters too faint to be seen.

A minority view is that the origin of the high redshifts, at least in some cases, is not due to the Hubble flow, and hence cannot be used in the standard way to determine the distance of the quasar. Rather, it is claimed that some component of the redshift is an intrinsic property of the quasar itself. Supporting evidence comes from the incidence of high redshift quasars near to low redshift bright galaxies — an unlikely state of affairs, advocates of the non-cosmological redshift would say, unless the quasars and bright galaxies are in some way related, by ejection of the quasars from the bright galaxies for example. A related problem is of quasar alignments (see *News and Views*, **289**; 533, 1981). An impressive number of pairs, and some multiple quasars have been found near to galaxies. The crucial question is whether these distributions are projection

effects, and hence occur by chance, or imply a genuine physical association. The majority of astronomers still doubt a physical association hypothesis because of the *a posteriori* statistics used, and also because no accepted theoretical framework exists to explain the large intrinsic redshifts. In order to calculate accurately the chances of finding a quasar near to a bright galaxy by projection effects, one needs to know the surface density of quasars. This number is very uncertain at faint magnitudes and different assumed densities will change radically the likelihood of a chance quasar/galaxy pair. So the question remains open, but it is unlikely that the majority view will change without evidence of physical association — as might be provided by an unambiguous detection of connecting material for example. Talks given during the rest of the meeting, addressed the question of what quasars are, and it was assumed without comment that quasars are distant and that redshift is a distance indicator.

The continuum radiation from a large number of quasars has now been observed from X-ray to radio frequencies. Variations in the intensity of this radiation over short timescales, in many cases days, strongly suggests an origin in the very central regions of the object. If quasars are assumed to be at great distances it is difficult to see what sort of powerhouse is capable of generating these enormous energies. A general picture which has gained acceptance over the last few years requires the presence, at the center of quasars and related objects, of a super-massive object (black hole) equivalent to something like 10^8 to 10^{10} solar masses. Material, possibly debris from stars torn apart by tidal forces, then falls onto the central object releasing energy as it does so. There is good evidence for a similar, although very much scaled down, version of

*The meeting was held in London on 13th February 1981 and was organised by M. G. Smith and R. F. Carswell.

Martin Ward is in the Institute of Astronomy, University of Cambridge.

process is the role played by soft X-rays which penetrate deep into the gas clouds. By considering this effect there is better agreement between the observed and the theoretically predicted emission line ratios. There is disagreement over whether obscuration by dust is important. This effect reduces the intensity of short wavelength emission lines relative to longer wavelength ones. It is certainly true in galaxies that where there's gas there's dust, but the relative amounts in quasar emission line clouds is at present unknown.

Finally the absorption line problem. It has been known for some time that the broad emission line Lyman alpha often has, superimposed on its profile, numerous narrow absorption lines which are usually blue-shifted with respect to the quasar. This is the so-called Lyman alpha 'forest'. The

question is, do these absorptions arise in material near to the quasar or elsewhere along the long line of sight between us and it? Opinion now favours the latter interpretation. An explanation in which the absorption occurs in the halos of galaxies is an interesting possibility. If this is correct we could use the light from background quasars to study the properties and statistics of galaxies far too distant to be detected by their own feeble light.

It would be wrong to give the impression that a revolution has taken place in our knowledge of what quasars are over the last few years, but neither is pessimism justified. The new multi-frequency continua and high resolution spectral line observations, together with variation studies have improved our understanding. Progress is being made. □

Triggering of volcanic eruptions by Earth tides

from R.S.J. Sparks

A recent study of Kilauea and Mauna Loa volcanoes in Hawaii (D. Dzurisin *Geophysical Research Letters* 7, 925; 1980) shows a weak but significant correlation of the Earth tide maxima with eruptions of Kilauea. No correlation at all was seen with the activity of Mauna Loa — a result which serves to emphasize how individual the behaviour pattern of very similar volcanoes can be.

The stress changes induced in the Earth's crust by fortnightly Earth tides are very small (maximum of 10^{-2} bar) compared with the levels of tectonic stress ($10^1 - 10^3$ bar) and the probable stresses required to fracture a volcanic edifice and trigger eruptions. Although many studies have attempted to link earthquakes with tidal stressing, the results have generally proved negative or inconclusive. One exception was a study of after-shocks by Ryall and others (*Bull. seism. Soc. Am.* 58, 215; 1968) where a good correlation with fortnightly Earth tides was demonstrated.

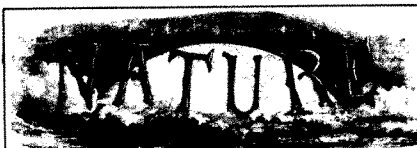
In a volcanic eruption, pressure builds up over periods of hours to months due to the ascent of lava into the volcano, the exsolution of volatiles or the entrapment of ground water. When the pressure on the volcanic edifice exceeds the strength of the rock, fracturing occurs and allows lava and gas to reach the surface. Fracture strengths of rock in tensile failure are probably in the range 10 to 100 bar, again very much larger than tidal stresses. Thus, at first sight, a correlation with Earth tides

seems unlikely given many other factors which will control the formation of suitable fractures.

Mauk and Johnston (*J. geophys. Res.* 78, 3356; 1973), however, have pointed out that it is the rate of stress accumulation that can be important. Fortnightly Earth tides can have stress accumulation rates of up to 10^{-5} bar s^{-1} . In volcanoes such as Kilauea the stress accumulation rates by increase in fluid pressure within the magma reservoir are as low as 2×10^{-7} bar s^{-1} . Thus, if stress is slowly building up Earth tides could trigger the eruption. On the other hand, stress accumulations of 5×10^{-5} to 8×10^{-4} bar s^{-1} , greatly exceeding the tidal rates, have been recorded from volcanoes in New Zealand.

The study by Dzurisin examines the two volcanoes, Kilauea and Mauna Loa, for which more data are available than any other volcano. The correlation between Earth tides and Kilauean eruptions is reasonably convincing — the probability of an eruption during a week centred on tidal maximum is nearly twice as much as the probability of an eruption in a week centred on tidal minimum. For Mauna Loa, however, there is no correlation. An important lesson is that it is often very dangerous to transfer the results from one volcano to another. Volcanologists should not use Earth tides in general predictions until they know the behaviour of their particular volcano very well. Most of the world's dangerous volcanoes, such as Mount St Helens, simply do not have a long enough record to find out whether there is a correlation. □

R.S.J. Sparks is in the Department of Earth Sciences, University of Cambridge.



100 years ago

THE ST. PETERSBURG DYNAMITE MINE

The following account of the bomb recently discovered in St. Petersburg, extracted from Russian sources, gives a remarkable picture of the state of society in the empire, where able chemists and expert miners can be found to engage in such desperate undertakings.

It appears from a sketch-plan which accompanied the translation put into our hand, that the mine extended from one side of Malaya Sadobaya Street to the centre of the roadway; the total length of the mine gallery being fifteen paces, the street must be thirty paces, say seventy-five feet wide.

The gallery terminated in a chamber about double its diameter, and in this was found the charge contained in a case twenty-two inches long and eight inches diameter, weighing sixty-five pounds, and beside this a glass jar contained about thirty pounds more of the explosive substance, apparently an excess quantity over that required for the actual explosion. The explosive consisted of a species of dynamite made by mixing nitro-glycerine with powdered charcoal. The description of the fuse, as contained in the Russian account, is very obscure, but would appear to have consisted of a wide heavy glass tube containing an explosive, described some time back in *Nature*, and prepared by mixing nitro-glycerine with about 10 per cent. of gun-cotton, the result being a very explosive substance of a partially gelatinous character. In the midst of this, and surrounded by a mixture of potassic chlorate and antimonious sulphide, was a sealed glass tube containing concentrated sulphuric acid and a leaden weight. The whole was then apparently connected with the dynamite in the case by means of an india-rubber tube also containing explosives. If this was the actual construction the *modus operandi* of the conspirators was very simple, for the heavy glass tube had only to be allowed to fall, when the lead would have broken the sulphuric acid tube, and the chlorate mixture would have at once inflamed, and the explosion of the jelly would have communicated by means of the rubber tube with the torpedo. At the same time it is very difficult to imagine the reasons which could have induced the conspirators to adopt so crude a method when, as it appears from the account, they had at their disposal in the room adjoining that from which the mine was driven no less than four galvanic batteries, with which the explosion could have easily been instantaneously effected by those on the watch for the passage of the intended victims, a method of operation much more consonant with the skilled character of the other parts of the work.

Whatever may have been the real mode intended to have been used by the conspirators, the results would have been sufficiently frightful, as it is probable that the charge found would have made a 'crater' of about fifty feet in diameter.

Some of our readers may be glad to learn that the French Association has resolved to curtail the number of scientific meetings in order to extend the time left for excursions and festivals. The programme includes two receptions by the Mayor of Algiers, one by the Governor-General, and a large Arabian *fête* called *Bitu*, dancing and singing by native women, &c. Possibly the reported massacre of the Flatters Expedition may put a stop to these ultra-scientific festivities.

From *Nature* 23, 7 April, 531, 542, 1881.

Erratum

In *Monitoring Mount St Helens* from M. McCormick (*News and Views* March 12, 290, 88, 1981) the value of aerosol mass added to the background stratosphere calculated from lidar data was incorrectly printed as being 0.4 to 0.6×10^{10} metric tons. The value should have been 0.4 to 0.6×10^6 metric tons (p.88, line 46) which agrees with that calculated from SAGE data. The predicted warming of the stratosphere near the peak aerosol concentration at 18km (p.88, line 62) was also incorrectly printed as 'a degree' instead of 'tenths of a degree'.

REVIEW ARTICLE

Membrane protein oligomeric structure and transport function

M. Klingenberg

Institut für Physikalische Biochemie der Universität München, Goethestrasse 33, 8000 München 2, FRG

Proteins which traverse membranes tend to have a dimeric structure in which the dimer is arranged asymmetrically across the membrane with the axis of symmetry perpendicular to the membrane plane. This general structure is well suited to the function of transporting nutrients across the cell membrane.

THERE is growing evidence that membrane proteins exist in oligomeric (usually dimeric) form, based on studies both of the intact membrane and of solubilized proteins. In reviewing the evidence here, I shall consider only those proteins which clearly or probably traverse the cell membrane. It is my aim to show that the structural features of these proteins have functional consequences which are compatible with the experimentally established prerequisites for carrier systems.

Examples of oligomeric membrane proteins

Studies of the intact membrane, for example with cross-linking reagents^{1,2}, suggest that some membrane proteins occur in oligomeric form but more compelling evidence has come from studies of the solubilized membrane proteins. Only in the past decade, by more judicious handling of nonionic detergents, has the isolation of membrane proteins reached a stage at which the purified proteins can confidently be considered undenatured and intact. This has allowed the oligomeric states of some solubilized membrane proteins to be demonstrated by direct measurements of molecular weight. Based on data from such studies I suggest that oligomericity can be used as a criterion to determine whether a membrane protein isolated in these detergents is intact, because cleavage of oligomers by detergents often leads to denaturation.

A survey of some selected membrane proteins with oligomeric structure is given in Table 1. Although much evidence has been derived from chemical cross-linking procedures, these are not free from artefacts. Less equivocal data can be obtained with the isolated protein, by determination of the molecular weight (by ultracentrifugation) and of the number of active sites per protein. Any complication due to artificial aggregation of hydrophobic proteins can be largely excluded by the appropriate use of nonionic detergents. If the functional molecular weight calculated from the number of sites per protein agrees with that determined by hydrodynamic methods, there can be little doubt that the isolated protein is in the native oligomeric state.

In the intact membrane, cross-linking experiments for detecting oligomeric proteins have rarely been successful³. Nevertheless, erythrocyte band 3 protein (identified with the anion carrier⁴) has been shown to occur as a dimer⁵ and to span the membrane⁶. Also, $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ has been cross-linked⁷ in the membrane to form a dimer of the large α and β subunits, which indicates that this pump may exist as a dimer $(\alpha\beta)_2$. The ADP/ATP carrier in mitochondrial membranes can be shown by combined cross-linking and immunochemical methods to be a dimer (M. L. Vitali and H. P. Schultheiss, unpublished data). For Ca^{2+} -ATPase there is some evidence of an oligomeric state in the membrane but it is equivocal⁸.

More abundant and clear-cut evidence for oligomeric structure is provided by the isolated membrane proteins. If solubilized with nonionic detergents, these proteins form micelles with an increased particle size due to the bound detergent. The detergent obscures the true hydrodynamic properties of the

proteins but this can be allowed for by use of appropriate corrections⁹. The membrane protein preparations in the nonionic detergents are mostly monodisperse and, according to several criteria, are undenatured, retaining their native structure. Using Triton X-100 my colleagues and I have isolated seven different membrane proteins from mitochondria¹⁰⁻¹³, chloroplasts¹⁴ and bacteria^{15,16} and have determined their molecular weights by ultracentrifugation¹⁷⁻¹⁹.

The band 3 protein of erythrocytes isolated with Triton X-100 is a dimer—this agrees with *in situ* cross-linking studies⁵. However, if solubilized by acetic acid, a dynamic monomer-oligomer equilibrium is observed²⁰. This preparation may not be native according to the criteria described above. Among the mitochondrial proteins isolated with Triton X-100, the ADP/ATP carrier was first recognized as a dimer with two identical subunits^{17,21}. Extending these methods, other mitochondrial proteins such as cytochrome *b* (ref. 10) and the cytochrome *bc*₁ complex^{13,18} were purified in defined oligomeric states. The highly hydrophobic cytochrome *b* was characterized as a dimer¹⁰ and the cytochrome *bc*₁ complex, consisting of several different subunits, was found to be a dimer with each subunit present at least twice^{18,22}. Other dimers isolated in Triton include the cytochrome *c* oxidase complex²³, the uncoupling protein from brown fat mitochondria¹⁹, the acetylcholine receptor from postsynaptic membranes and the triose phosphate- P_i exchange carrier of chloroplasts¹⁴. Two membrane-bound dehydrogenases from bacteria, formate dehydrogenase¹⁵ and fumarate reductase¹⁶, which also form dimers, are not listed in Table 1 because they do not penetrate the cell membrane and bind only a relatively small amount of Triton.

When isolated with the nonionic detergent Lubrol, $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ is a dimer by ultracentrifugal studies^{24,25}; in the same detergent Ca^{2+} -ATPase is a tetramer⁸ but this may result from nonspecific aggregation as the tetramer can be dissociated into monomers by greater amounts of nonionic detergent

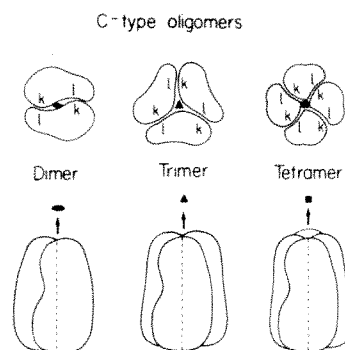


Fig. 1 Protein oligomers with n -fold rotational symmetry axis consisting of 2, 3 and 4 identical subunits. The repeating interfaces between the subunits are designated k and l .

Table 1 Some oligomeric transmembrane proteins

Protein	Source	Molecular weight ($\times 10^3$)		Type	Evidence (ref.)
		Monomer	Oligomer		
ADP/ATP carrier	Mitochondria: animal and <i>Neurospora crassa</i>	30 34	60 68	α_2 α_2	Inhibitor binding (21, 69) Cross-linking (17, 69) Hydrodynamic (17)
Uncoupling protein	Brown adipose tissue (mammalian)	32	64	α_2	Hydrodynamic (19)
($\text{Na}^+ + \text{K}^+$)ATPase	Plasma membrane, for example, shark rectal gland	—	330	$(\alpha\beta)_2$ $(\alpha\beta_2)_2$	Cross-linking (7) Hydrodynamic (24, 25)
Ca^{2+} -ATPase	Sarcoplasmic reticulum	100	400	α_4	
Anion carrier (band III)	Erythrocytes	90	180	α_2	Cross-linking (5)
Glucose carrier	Erythrocytes	55	100	α_2	Freeze-etching counts (33) Inhibitor binding (33, 70)
Galactoside carrier	<i>E. coli</i>	30, 44?	—	α_2	Cited in ref. 71
Porin	<i>E. coli</i> , outer membrane	37	—	α_3	Electron diffraction (72)
Bacteriorhodopsin	<i>Halobacterium</i>	28	—	(α_3)	Electron diffraction (28)
Acetylcholine receptor	Electric organ: <i>Torpedo</i>	250	500	$(\alpha\beta \dots)_2$	Hydrodynamic (73) Disulphide formation (74)
Cytochrome <i>b</i>	Mitochondria: animal and <i>N. Crassa</i>	30	60	α_2	Hydrodynamic (10)
Cytochrome <i>bc_1</i> -complex	Mitochondria	—	420	$(\alpha_2\beta \dots)_2$	Hydrodynamic (18, 22)
Cytochrome oxidase	Mitochondria	110	345	$(\alpha\beta \dots)_2$	Hydrodynamic (23)

without loss of ATPase activity²⁶. In contrast, rhodopsin occurs as a monomer both in the membranous and isolated states²⁷. The diffusibility of this protein as part of its function is in agreement with its monomeric state. Bacteriorhodopsin, which forms a crystalline purple membrane lattice²⁸, is also monomeric in detergent solution²⁹, although electron diffraction indicates trimer formation in the membrane.

Another fact which strongly supports dimeric structures comes from ligand binding studies, using highly specific inhibitors of translocators. The frequent result of such experiments is that only half a binding site exists per protein subunit: that is, a complete binding site requires at least a dimer of the protein subunits. Examples are the binding of ouabain to ($\text{Na}^+ + \text{K}^+$)ATPase³⁰ (see however ref. 31), the binding of atractylate or bongkrekate to ADP/ATP carrier^{21,32} and cytochalasin B binding to the glucose carrier³³.

There is evidence that the dimeric structure in the isolated protein is the native one for the ADP/ATP carrier as it loses ligand binding affinity on forming monomers. Only in the dimeric form does the cytochrome *bc_1* complex show succinate-cytochrome *c* reductase activity²². The ($\text{Na}^+ + \text{K}^+$)ATPase is reconstitutionally active when isolated with nonionic detergent as a dimer³⁴.

To conclude, note that some oligomeric membrane proteins have a rather cryptic transport function for ions: for example, cytochrome complexes may pump H^+ (ref. 35), various ATPases pump H^+ , Na^+ or Ca^{2+} , the acetylcholine receptor has a Na^+ channel³⁶, and the uncoupling protein of mitochondria catalyses anion transport³⁷.

Structural principles of oligomeric membrane proteins

Further analysis of the implications and characteristics of the oligomeric structure of membrane proteins requires a brief resumé of some general principles of oligomer formation. Much has been published on the symmetry properties of oligomeric enzymes and their functional consequences³⁸⁻⁴⁰, therefore these problems will be discussed only with regard to the particular situation of membrane proteins. Figure 1 shows some oligomers with 'C'-type symmetry. A closer look at the geometrical principle of association is important for understanding the characteristics of these structures. Dimeric and trimeric proteins can form only one type of symmetrical oligomer which is characterized by a C_2 - and C_3 -rotational axis. Tetramers can form several symmetrical structures; the oligomer with C_4 symmetry corresponds to those of the dimers and trimers. Higher symmetry of the D_2 type is found in many soluble tetrameric proteins. For reasons discussed below, this arrangement is

unlikely for tetrameric membrane proteins. Tetramers can also be visualized as being arranged in pseudo- C_4 symmetry and strict C_2 symmetry when two dimers associate to form a tetramer by retaining their original dimeric structure.

One point of discussion concerning oligomeric structure is that of isologous versus heterologous association and the relationship of this to cooperative effects between the subunits^{41,42}. Regarding this differentiation, only the dimer would be isologous whereas the C_3 -trimer and the C_4 -tetramer are heterologous. This differentiation, however, can be dismissed in analysing the contact area between the subunits in these oligomers. Each subunit may be thought to have a k- and an l-surface separated by the symmetry axis; in all three cases the k-surface associates with the l-surface of the other subunit in an 'heterologous' manner. This is true for the dimer in much the same way as for the trimer and tetramer. There is no particular distinction in the 'isologous' association of the dimer over those of the trimer and tetramer. In fact, one of the main advantages claimed for isologous as compared with heterologous association, which applies to all three cases, is that a mutation which changes the interface will be effective on both sides of the C-axis in the oligomer.

Oligomer membrane proteins and phospholipid bilayer

Membrane proteins are generally assumed to have a hydrophobic surface in those areas which make contact with the membrane lipid layer. The share of the hydrophobic surface should be particularly large in proteins deeply imbedded in the membrane, the most prominent example of which are carriers

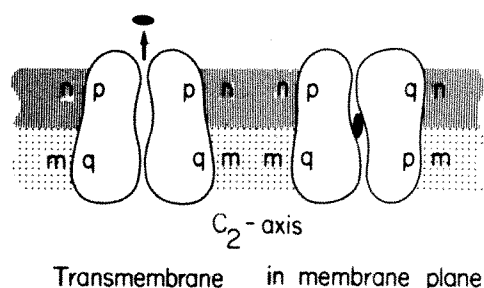


Fig. 2 Interaction of a dimeric protein with the asymmetric phospholipid bilayer. The localization of a dimer with the 2-fold axis either normal or parallel to the membrane plane. The surfaces of the protein are designated by p and q and the phospholipid bilayers by m and n.

such as those for ADP and ATP or glucose. The high amount of Triton binding to these proteins has been thought to reflect the extensive membrane contact area⁴³. However, the amount of detergent bound cannot be used as a measure of the hydrophobic region because of the tendency of nonionic detergents to form micelles at these surfaces⁴⁴.

The 'specific' interaction between subunits should be contrasted with nonspecific association often encountered in membrane proteins resulting from the interaction of the hydrophobic surfaces. This aggregation can occur if the isolated protein lacks sufficient detergent. The size of the aggregation may reach microscopic dimensions or it may be limited and even well defined, for example, in protein micelles with a hydrophobic core.

How are oligomeric membrane proteins arranged in the membrane? Biomembranes are topologically asymmetric for the obvious reason that they separate spaces of different composition. There is a prominent asymmetry between both leaflets of the lipid bilayer caused by asymmetric distribution of membrane phospholipid⁴⁵. The asymmetric assembly of the membrane can also include the membrane proteins. For the oligomeric transmembrane protein an arrangement with the rotational axis normal to the membrane is the obvious choice. This arrangement might be stabilized by interaction with the asymmetric bilayer, particularly by the different phospholipid head groups which can interact with the membrane protein through ionic forces.

Due to the interaction between phospholipids and protein, a transmembrane axial arrangement is energetically more favourable than the symmetrical arrangement with the axis within the membrane plane. This can be illustrated (Fig. 2) by summing up the free energy changes associated with the interaction of specific protein surfaces and phospholipid regions of the asymmetric bilayer. The free energy change for the transmembrane axial case (ΔG_T) can be described as being equivalent to $(np + mq) \times 2$ (see Fig. 2 legend for explanation of symbols) whereas the free energy change for the parallel axis would be $\Delta G_P = np + mq + nq + mp$. Assuming that $np + qm$ is the preferred interaction (a free energy minimum), for example, $np + qm > nq + mp$, it follows that $\Delta G_T > \Delta G_P$. These arguments are valid only for oligomeric proteins and we can conclude that the asymmetric transmembrane axial arrangement provides the membrane proteins not only with maximum interaction forces with phospholipids but also with additional stabilization of their quaternary structure.

After solubilization of membrane proteins in detergent, the asymmetric environment is abolished. By replacing the phospholipid bilayer, nonionic detergents form a micelle around the protein⁴⁴. In general the detergent will be asymmetrically distributed around the protein, but the uniform detergent does not allow a differential binding of molecules to different parts of the protein. The absence of this differentiation may contribute to the relative instability of detergent-solubilized proteins.

Membrane multicomponent complexes

A clear example of a C_2 -transmembrane arrangement of a dimeric membrane protein is the $(Na^+ + K^+)$ -pump (see also ref. 46). The sections responsible for ATP hydrolysis in both subunits must be directed towards the cytosolic face and electron microscopy shows that $(Na^+ + K^+)$ ATPase (and Ca^{2+} -ATPase⁴⁷) bulge only towards the inner side of the cell. Each large subunit of the $(Na^+ + K^+)$ ATPase binds a minimum of one 53,000 molecular weight (M_r) glycoprotein which faces the extracellular space^{31,48}.

In C_2 -transmembrane multicomponent proteins the subunits can be expected to be attached asymmetrically to the membrane, because identical intersubunit binding sites follow the transmembrane axis. Thus, in membrane-bound multi-enzyme complexes, subunits will be aligned in such a way that they form a C_2 -symmetric complex. Often only one or two of the peptides are imbedded in the membrane, the other subunits being attached to the surfaces protruding from the membrane,

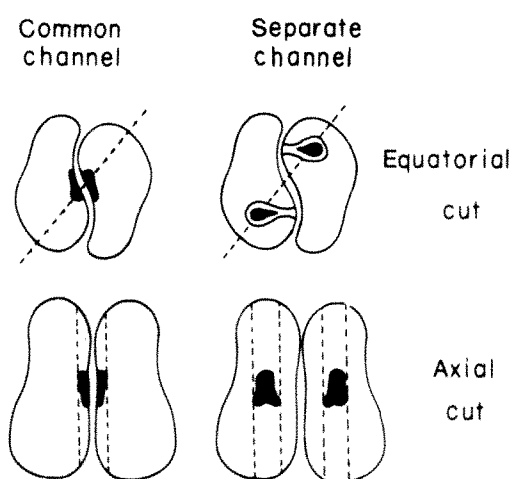


Fig. 3 Common central or separate channels and the binding centres in dimeric transport proteins.

for example in the ATP-synthase of mitochondria. Conversely, the existence of detachable subunits on one side of a membrane protein can be considered as evidence for C -type symmetry. For oligomeric membrane-bound glycoproteins, a transmembrane asymmetry is obligatory as the glycoside moiety is always directed towards the outer surface of the organelle membrane. This is the case, for example, with the anion channel of erythrocytes which is a dimer of two identical transmembrane glycoprotein subunits^{48,49}. Glycophorin may also form a transmembrane dimer in the erythrocyte, although no clear evidence has been reported⁵⁰.

Oligomeric structure and transport

So far we have seen that membrane proteins may acquire certain structural advantages unavailable to soluble proteins by forming oligomers in their hydrophobic environment. Do they also derive functional benefits from the oligomeric, axial transmembrane arrangement?

Translocation of hydrophobic molecules through the membrane requires a hydrophilic channel. Relatively rigid channels extending through the membrane exist in permanent forms to allow, for example, the nonspecific passage of solutes. Closely related to such pores are the gated channels through which, when open, many molecules or ions may diffuse. For the usual carrier-mediated transport, however, the channel should close or open during each turnover when one molecule or a defined set of ions is being moved through the membrane. In this case, the channels are highly flexible and gated so that at no instance do they freely extend across the whole membrane.

Binding forces between subunits are at a minimum along the symmetry axis because identical amino acid residues oppose each other. This should facilitate formation of a straight channel along the symmetry axis and corresponding pores have indeed been observed at atomic resolution for dimeric and tetrameric enzymes. With a transmembrane arrangement of the symmetry axis, a channel across the membrane could be formed in an oligomeric membrane protein (Fig. 3). Furthermore, this channel might have a hydrophilic surface as in membrane proteins a hydrophilic subunit interface may be energetically more favoured in the formation of oligomers, in view of the hydrophobic outer surface.

It has been tempting to suppose that membrane-bound carriers use the intersubunit channel as a translocation path for the substrate^{46,51} but experimental evidence in favour of such a model has only recently been obtained. The gated pore model was suggested from studies on the dimeric ADP/ATP carrier⁵², as will be discussed below. The alternative view has been that each subunit forms a separate channel. These channels could run either inside each subunit or at the interface (Fig. 3). However,

outside the symmetry axis, *a priori*, no straight path can be expected to occur.

Experimental data from exchange type systems⁵² suggest an intersubunit translocation path. (This is rather surprising because *a priori*, a reciprocating or 'flip-flop' mechanism of the two subunits would be the more obvious.) Two cooperating parallel channels are required for the simultaneous release and uptake of the exchange molecules. A dimeric carrier seems to be an obvious choice for this although alternatively the exchange could function in a sequential manner, often denoted as a 'ping-pong' mechanism. Each channel operates by facilitating, in turn, the uptake and release of a substrate molecule. The ping-pong performance is controlled by the gates in the channel. This exchange system differs from the unidirectional transport in that the gate can function only when the substrate is present. Either the common channel or the separate channels in a dimer could function by this process. Experimental evidence is largely in favour of this mechanism as opposed to the reciprocating model^{53,54}.

Binding sites and gates of the carrier

The translocation apparatus requires, besides the channel, a binding site for the substrate and a closely linked gate. It is the study of the binding sites that has so far produced the most valuable information concerning carrier mechanisms, for it is the binding sites of carriers which determine the high specificity of transport. It becomes abundantly clear from specificity studies that the binding of the transported molecule immediately triggers the transport reactions, presumably by means of conformation changes which open the channel gates. Substrate analogues or 'inhibitors' may bind with even higher affinity but have little or no triggering power.

In dimeric carriers, the two binding sites may be separated (Fig. 3), in which case the separate translocation paths would be imbedded within the subunits. With a central channel, the binding sites would directly face each other at the C_2 -axis and as a result, binding of one ligand should prevent that of the other. In fact, as shown above, in several dimeric carriers, binding of only one substrate or inhibitor per dimer has been observed.

The behaviour of some dimeric or tetrameric enzymes, in which negative cooperativity between the subunits controls binding and catalytic activity^{55,56}, has been called 'half-site' reactivity^{40,57}. Is the observed 'half-site' binding of carriers due to a similar effect? If so, little would be gained for dimeric carriers with separate channels, except for a decrease in transport activity. However, for the counter-exchange, 'half-site' reactivity would have a purpose if it generated a reciprocating

transport, which would be analogous to the flip-flop activity elaborated in detail for some half-site enzymes⁵⁷. The two binding sites are directed to opposite membrane sides and thus alternate between binding and releasing substrate to opposite sides. A flip-flop model has been discussed in detail for the $(Na^+ + K^+)ATPase$ ⁵⁸. Most of the evidence for $(Na^+ + K^+)ATPase$, however, is against counterflow and in favour of a sequential $(Na^+ + K^+)$ -exchange.

In the central channel model half-site reactivity becomes a trivial sterical consequence as the two binding sites face each other. Binding of one substrate molecule would sterically prevent binding to the second site, particularly if the substrate is relatively large. As a result, functional asymmetry between both subunits is induced, which may further generate some conformational asymmetry.

The translocation process

The actual translocation process can be accommodated by assuming an alternating gating of the channel around the binding site such that the binding site is opened either to the internal surface of the membrane, when the carrier is said to be in the 'i' state, or to the external surface, when the carrier is in the 'e' state. In the case of an axial binding centre, the states are distinguished by a positional shift between the two subunits such that gates are closed or opened at the interface. With separate channels, a conformational change inside each subunit during the transition would be required and the gating can be thought of as being caused by positional shifts between the protein lobes around the channel or by discrete shifts of only one or two amino acid residues.

Evidence at the molecular level for the two translocational states of the carrier was first deduced by studying the binding site of the ADP/ATP carrier⁵². The transition between the 'e' and 'i' state (in that instance, 'c' = cytosolic state and 'm' = matrix state) had a strict requirement for substrate binding and as predicted, the binding centre was shown never to face both sides of the membrane simultaneously. These conclusions were fully corroborated with the solubilized, purified ADP/ATP carrier^{53,59}.

The radical change in specificity for the inhibitors of the binding centre on the transition between the 'c' and 'm' state has been interpreted, as illustrated in Fig. 5, to reflect the existence of a fixed rather than a rotational carrier⁵⁴. With the former an asymmetric substrate should, on entering the binding centre from one side, interact with the protein through a different face than on leaving from the opposite side. By contrast, with a rotational carrier the relative position of substrate to carrier should not be changed; in a stationary carrier the binding surface of the protein state should differ from that in the 'e' state so that there is a drastic difference in specificity. That is precisely what is seen with the ADP/ATP carrier which specifically binds atracylate in the 'c' state but bongkredate in the 'm' state, whereas the affinity for the substrates remains largely unchanged⁴⁷. Similarly, the glucose carrier in erythrocytes binds cytochalasin B only from inside⁶⁰ and binds various steroids asymmetrically from inside and out⁶¹; the Na^+ -dependent glucose carrier has a different ordered binding sequence for Na^+ and glucose on the sides⁶²; and the $(Na^+ + K^+)ATPase$ exhibits affinity changes for Na^+ and K^+ in the E_1 and E_2 state (although here the superimposition of an ATP-driven transition complicates the same principle).

The model of Fig. 5 requires that a single binding site changes its specificity between the 'i' and 'e' states of the carrier. Alternatively there are two different binding sites in series along the channel, either axial or intrasubunit, with the substrate migrating between the two sites (Fig. 4). The different specificity would then result from the dissimilar configuration of the 'i' and 'e' site located in dissimilar regions of protein. However, from a comparative viewpoint this seems improbable because enzymes generally contain only one binding centre per subunit, with second centres being regulatory.

A configurational change in a binding site between the 'i' and 'e' state implies that a partially different set of amino acid

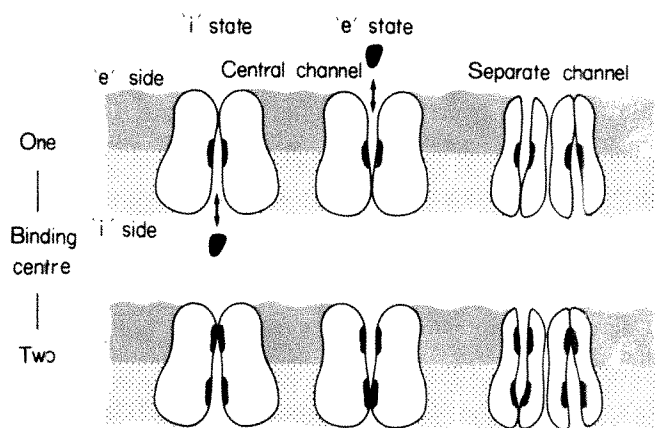


Fig. 4 Alternative 'gated pore' models of membrane carriers. With a single central channel the binding centre opens either to the outside (e-side) or inside (i-side) primarily by repositioning of the two subunits. For separate channels the gating requires intrasubunit conformational change.

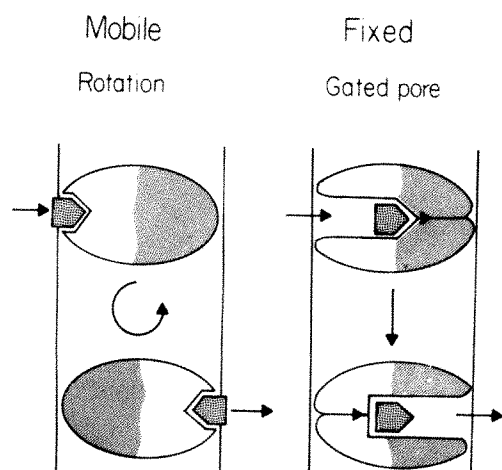


Fig. 5 The difference in flexibility of the binding centre in a rotational and fixed carrier⁵².

residues participates in the binding in the two states; the consequential change of affinity for the substrates could be beneficial in unidirectional transport. On the other hand, in counter-exchange, the binding affinities for the substrate would need to be well balanced, with energy-driven transport perhaps being achieved by a superimposed imbalance of the binding affinities in the two states. Evidence for the ADP/ATP carrier indicates that energy, as exerted by a membrane potential, does not change affinities but rather the distribution between the 'i' and 'e' states^{63,64}.

In trying to distinguish between the various models for translocation shown in Fig. 4, perhaps the best evidence for several dimeric carriers is that they bind only one molecule of substrate or inhibitor. Although this is still compatible with the existence of two separate channels in which the second site is inhibited by negative cooperativity, it seems unlikely because it has not been possible to increase binding at the ADP/ATP carrier by abolishing negative cooperativity, for example by monomerization of the carrier.

Another important observation is the conformational change on transition between the 'i' and 'e' state. This is most compatible with the central channel model where both subunits change conformation between the 'i' and 'e' state in parallel. More complicated assumptions must be made for the two-channel

model: conflicting cooperativity is required, a negative one to lock the other binding centre, and a positive one to synchronize the conformational transition between the binding and non-binding subunits. Furthermore, the reciprocating (flip-flop) model can be discounted because it requires that one subunit always exists in the 'i' state and the other in the 'e' state.

Most of these constraints are abolished if there are independent separate channels. In that case both binding sites of dimeric carriers should be active. Evidence for that exists for the anion carrier of erythrocytes⁶⁵ and the galactoside carrier of *Escherichia coli*⁶⁶ but has yet to be confirmed for the purified native protein.

Dimeric membrane proteins with other than transport functions, such as cytochrome complexes in mitochondria, also seem to show half-site reactivity. For example, cytochrome *b*, being a dimer in the cytochrome *bc*₁ complex, reacts as if there are two different species of cytochrome *b*. An anti-cooperativity between two identical cytochrome *b* molecules has been invoked for an alternating H⁺-pumping activity of cytochrome *b* in the redox cycle⁶⁷. A similar suggestion has been derived by Wikström⁶⁸ for the cytochrome oxidase complex. Note, however, that another H⁺-pump, bacteriorhodopsin, seems to operate with independent channels in each subunit. This can be reconciled with the fact that interaction between bacteriorhodopsin trimers in the hexagonal crystalline arrangement of the purple membrane is rather weak.

Conclusions

Oligomeric structures seem to be particularly well suited to integral membrane proteins which have carrier functions. The obvious role in transport of many integral proteins requires channels extending through the protein either along the *n*-fold axis or inside each subunit. The channels may be gated during each substrate turnover as in typical carrier-mediated transport. The catalytic binding centre in the channel of a number of dimeric carriers exhibits half-site reactivity which can be interpreted either as negative cooperativity for separate channels or as steric overlap in a common axial channel. Gating, as triggered by substrate binding, opens the site to the inside and outside, in turn. As a result the specificity of the carrier to inhibitor ligands is drastically changed in accordance with the direction in which transport is proceeding, for a stationary carrier. The parallel conformational change of both subunits between the 'e' and 'i' state, together with the half-site binding, can best be explained in terms of a single centre axial channel.

- Peters, K. & Richards, F. M. A. *Rev. Biochem.* **46**, 523-551 (1977).
- Das, M. & Fox, C. F. A. *Rev. Biophys. Bioengng.* **8**, 165-193 (1979).
- Wang, K. & Richards, F. M. J. *biol. Chem.* **249**, 8005-8018 (1974).
- Cabantchik, Z. O., Knauf, P. A. & Rothstein, A. *Biochim. biophys. Acta* **515**, 239-302 (1978).
- Yu, J. & Steck, T. L. *J. biol. Chem.* **250**, 9176-9184 (1975).
- Bretscher, M. S. *J. molec. Biol.* **59**, 351-356 (1971).
- Kyte, J. J. *biol. Chem.* **260**, 7443-7448 (1975).
- LeMaire, M., Jørgensen, K. E., Røigard-Petersen, H. & Møller, J. V. *Biochemistry* **15**, 5805-5812 (1976).
- Tanford, C., Nozaki, Y., Reynolds, J. A. & Makino, S. *Biochemistry* **13**, 2369-2376 (1974).
- von Jagow, G. *et al.* *FEBS Lett.* **91**, 121-125 (1978).
- Klingenberg, M., Riccio, P. & Aquila, H. *Biochim. biophys. Acta* **503**, 193-210 (1978).
- Lin, C. S. & Klingenberg, M. *FEBS Lett.* **113**, 209-213 (1980).
- Riccio, P., Schagger, H., Engel, W. D. & von Jagow, G. *Biochim. biophys. Acta* **459**, 250-262 (1977).
- Flügge, U. I. & Heldt, H. W. in *Function and Molecular Aspects of Biomembrane Transport* (eds Quagliariello, E. *et al.*) 373-382 (Elsevier, Amsterdam, 1979).
- Kröger, A., Winkler, E., Innerhofer, A., Hackenberg, H. & Schagger, H. *Eur. J. Biochem.* **94**, 465-475 (1979).
- Uden, G., Hackenberg, H. & Kröger, A. *Biochim. biophys. Acta* **591**, 275-288 (1980).
- Klingenberg, H. & Klingenberg, M. *FEBS Lett.* **56**, 113-138 (1975).
- von Jagow, G., Schagger, H., Riccio, P., Klingenberg, M. & Kolb, H. J. *Biochim. biophys. Acta* **462**, 549-559 (1977).
- Lin, C. S., Hackenberg, H. & Klingenberg, M. *FEBS Lett.* **113**, 214-216 (1980).
- Dorst, H. J. & Schubert, D. *Hoppe-Seyler's Z. physiol. Chem.* **360**, 1605-1618 (1979).
- Riccio, P., Aquila, H. & Klingenberg, M. *FEBS Lett.* **56**, 113-138 (1975).
- Weiss, H. & Kolb, H. J. *Eur. J. Biochem.* **99**, 139-149 (1979).
- Robinson, N. C. & Capaldi, R. A. *Biochemistry* **16**, 375-380 (1977).
- Hastings, D. F. & Reynolds, J. A. *Biochemistry* **18**, 817-821 (1979).
- Esmann, M., Christiansen, C., Karlsson, K. A., Hansson, G. C. & Skou, J. C. *Biochim. biophys. Acta* **603**, 1-12 (1980).
- Dean, W. L. & Tanford, C. *Biochemistry* **17**, 1683-1690 (1976).
- Osborne, H. B., Sardet, C. & Hellenius, A. *Eur. J. Biochem.* **44**, 383-390 (1974).
- Henderson, R. A. *Rev. Biophys. Bioengng.* **6**, 87-109 (1977).
- Reynolds, J. A. & Stoekenius, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2803-2804 (1977).
- Jørgensen, P. L. in *Function and Molecular Aspects of Biomembrane Transport* (eds Quagliariello, E. *et al.*) 81-90 (Elsevier, Amsterdam, 1979).
- Craig, W. S. & Kyte, J. J. *biol. Chem.* **255**, 6262-6269 (1980).
- Aquila, H., Eiermann, W., Babel, W. & Klingenberg, M. *Eur. J. Biochem.* **85**, 549-560 (1978).
- Hinkle, P. C., Sogin, D. C., Wheeler, T. J. & Telford, J. N. in *Function and Molecular Aspects of Biomembrane Transport* (eds Quagliariello, E. *et al.*) 487-494 (Elsevier, Amsterdam, 1979).
- Hilden, S. & Hokin, L. E. *J. biol. Chem.* **250**, 6296-6303 (1975).
- Wikström, M. K. F. & Krab, K. in *Proc. 29th Mosbach Colloq. Energy Conservation in Biol. Membranes* (eds Schäfer, G. & Klingenberg, M.) 128-139 (Springer, Berlin, 1978).
- Heidmann, T. & Changeux, J. P. A. *Rev. Biochem.* **47**, 317-357 (1978).
- Nicholls, D. G. *Biochim. biophys. Acta* **549**, 1-29 (1979).
- Cornish-Boden, A. J. & Koshland, D. E. Jr *J. biol. Chem.* **246**, 3092-3102 (1971).
- Matthews, B. W. & Bernard, S. A. A. *Rev. Biophys. Bioengng.* **2**, 257-312 (1973).
- Seydoux, F., Malhotra, O. P. & Bernhard, S. A. *CRC Crit. Rev. Biochem.* **2**, 227-257 (1974).
- Monod, J., Wyman, J. & Changeux, J. P. *J. molec. Biol.* **12**, 88-118 (1965).
- Hanson, K. R. *J. molec. Biol.* **22**, 405-409 (1966).
- Hellenius, A. & Simons, K. *Biochim. biophys. Acta* **415**, 29-79 (1975).
- Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membrane* 2nd edn, 209 (Wiley, New York, 1980).
- Op den Kamp, J. A. F. A. *Rev. Biochem.* **48**, 47-71 (1979).
- Singer, S. J. A. *Rev. Biochem.* **43**, 805-833 (1974).
- MacLennan, D. H. & Holland, P. C. in *Enzymes of Biological Membranes: Membrane Transport* Vol. 3 (ed. Martonosi, A. N.) 221-259 (Plenum, New York, 1976).
- Jenkins, R. E. & Tanner, M. J. A. *Biochem. J.* **161**, 139-147 (1977).
- Steck, T. L. & Dawson, G. *J. biol. Chem.* **249**, 2135-2142 (1974).
- Marchesi, V. T., Furthmayr, H. & Tomita, M. A. *Rev. Biochem.* **45**, 667-698 (1976).
- Jardetzky, O. *Nature* **211**, 969-970 (1966).
- Klingenberg, M. in *Enzymes of Biological Membranes: Membrane Transport* Vol. 3 (ed. Martonosi, A. N.) 383-438 (Plenum, New York, 1976).
- Klingenberg, M., Aquila, H., Krämer, R., Babel, W. & Feckl, J. in *Biochemistry of Membrane Transport* (eds Semenza, G. & Carafoli, E.) 567-579 (Springer, Heidelberg, 1976).

54. Klingenberg, M., Riccio, P., Aquila, H., Buchanan, B. B. & Grebe, K. in *The Structural Basis of Membrane Function* (eds Hatefi, Y. & Djavadi-Ohanian, E.) 293–311 (Academic, New York, 1976).
55. Stallcup, W. B. & Koshland, D. E. Jr *J. molec. Biol.* **80**, 41–62 (1973).
56. Levitzki, A. & Koshland, D. E. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1121–1128 (1969).
57. Lazdunski, M. in *Current Topics in Cellular Regulations* (eds Horecker, B. & Stadtman, E.) 267–309 (Academic, New York, 1972).
58. Repke, K. R. H. et al. *Ann. N.Y. Acad. Sci.* **242**, 203–219 (1974).
59. Aquila, H. & Klingenberg, M. in *Function and Molecular Aspects of Biomembrane Transport* (eds Quagliariello, E. et al.) 291–303 (Elsevier, Amsterdam, 1979).
60. Deves, R. & Krupka, R. M. *Biochim. biophys. Acta* **510**, 339–348 (1978).
61. Krupka, R. M. & Deves, R. *Biochim. biophys. Acta* **598**, 134–144 (1980).
62. Hopfer, U. & Groseclose, R. J. *biol. Chem.* **255**, 4453–4462 (1980).
63. Krämer, R. & Klingenberg, M. in *Function and Molecular Aspects of Biomembrane Transport* (eds Quagliariello, E. et al.) 515–518 (Elsevier, Amsterdam, 1979).
64. Klingenberg, M. in *Mitochondria/Biomembranes, Proc. 8th FEBS Meet.*, Amsterdam, 147–162 (Elsevier, Amsterdam, 1972).
65. Jennings, M. L. & Passow, H. *Biochim. biophys. Acta* **554**, 498–519 (1979).
66. Overath, P., Teather, R. M., Simoni, R. D., Aichele, G. & Wilhelm, U. *Biochemistry* **18**, 1–11 (1979).
67. van Jagow, G. & Engel, W. D. *FEBS Lett.* **111**, 1–5 (1980).
68. Wikström, M. in *The Proton and Calcium Pumps* (eds Azzzone, G. F. et al.) 215–226 (Elsevier, Amsterdam, 1978).
69. Klingenberg, M., Hackenberg, H., Eisenreich, G. & Mayer, I. in *Function and Molecular Aspects of Biomembrane Transport* (eds Quagliariello, E. et al.) 291–303 (Elsevier, Amsterdam, 1979).
70. Baldwin, S. A., Baldwin, J. M., Gorga, F. R. & Lienhard, G. E. *Biochim. biophys. Acta* **552**, 183–188 (1979).
71. Büchel, D. E., Gronenborn, B. & Müller-Hill, B. *Nature* **283**, 541–545 (1980).
72. Rosenbusch, J. P., Steven, Y. C., Alkan, M. & Regenass, M. in *Regular Two-Dimensional Arrays in Biomembranes* (ed. Baumeister, W.) (Springer, Berlin, in the press).
73. Reynolds, J. A. & Karlin, A. *Biochemistry* **17**, 2035–2038 (1978).
74. Hamilton, S. L., McLaughlin, M. & Karlin, A. *Biochemistry* **18**, 155–163 (1979).

Editor's Note: An article by Dr Jack Kyte of the University of California, San Diego, in which are presented some alternative views on the mechanistic role of membrane proteins in active transport, will be published shortly in *Nature*.

ARTICLES

The 'braided' F-ring of Saturn

Stanley F. Dermott

Center for Radiophysics and Space Research, Space Sciences Building, Cornell University, Ithaca, New York 14853, USA

The larger of the two satellites which bound the F-ring of Saturn could have only one first-order resonance within the ring. If such a resonance exists, then a wave pattern of equally spaced loops which co-rotate with the satellite will be excited. This pattern could account for the braided appearance of the ring.

GOLDREICH AND TREMAINE¹ consider that particles in the narrow rings of Uranus are confined by tidal torques exerted on their orbits by small, unseen, satellites. The recent discovery by Voyager 1 of two small satellites, S13 and S14, which bound the narrow F-ring of Saturn², seems to confirm this view. However, the structure of the F-ring is complex and quite unlike that of the ϵ ring of Uranus. Some of the images of the F-ring show two bright rings which appear to be intertwined or braided. These two rings are closely associated with a third, but this is very diffuse (see Fig. 1)³. We show here how gravitational perturbation of the ring particle orbits by S13 and S14 may not only account for the confinement of the ring particles but may also explain some of the 'braided' structure of the F-ring.

Tidal torques

The tidal torque exerted on a narrow ring of particles of total mass m_r by a nearby satellite of mass m can be estimated from the scattering angle δ in Fig. 2b (ref. 4).

$$\delta = \frac{2Gm}{xU^2} \quad (1)$$

where the relative velocity U is given by

$$U = \pm \frac{3}{2}nx \quad (2)$$

G is the gravitational constant, n is the mean motion of the ring particles and x is the separation of the satellite and the ring. Since the tangential component of the relative velocity is reduced from U to $U \cos \delta$, angular momentum is exchanged with the net result that the ring experiences a torque

$$T = \pm \frac{4}{9\pi} \left(\frac{Gm}{nx^2} \right)^2 m_r \quad (3)$$

The sign of the torque depends on whether the satellite is on the inside or the outside of the ring and is such that the satellites repel the ring particles¹.

If a ring is bounded by two satellites, then its width will be reduced until the confining torques just counteract the tendency

of the ring to spread due to particle collisions. The equilibrium width of the ring is

$$W \sim \left(\frac{3\pi\tau}{4} \right)^{1/2} \left(\frac{x}{r} \right)^{5/2} \left(\frac{M}{m} \right) d \quad (4)$$

where M is the mass of the planet, r is the radius of the ring and d is the characteristic diameter of those particles which determine the spreading rate¹. τ is not simply the optical depth of the ring, this is always determined by the number density of the smaller particles, but the optical depth that the ring would have if all particles of diameter $< d$ were removed. For the F-ring of Saturn

$$d \sim \frac{40}{\tau^{1/2}} \left(\frac{30 \text{ km}}{W} \right) \left(\frac{m/M}{9 \times 10^{-9}} \right) \left(\frac{1,000 \text{ km}}{x} \right)^{5/2} m \quad (5)$$

($m/M \sim 9 \times 10^{-9}$ corresponds to a satellite of diameter 200 km and density 1.2 g cm^{-3} .) If Goldreich and Tremaine are right, then the F-ring must have a substantial population of large particles. However, we consider that equation (3) may overestimate the magnitudes of the confining torques.

When the ring is in equilibrium, the torques due to the satellites must be equal and opposite, hence, from equation (3),

$$\frac{m_o}{x_o^2} = \frac{m_i}{x_i^2} \quad (6)$$

where subscripts o and i refer to the outer and inner satellite respectively. For the purposes of this article, it is sufficient to assume that the satellites and the ring particles move in coplanar, circular orbits.

Standing waves

Because $x \ll r$, the gravitational force on a ring particle due to a satellite is only effective at close encounters. During encounter, a particle initially moving in a circular orbit, acquires a radial velocity $U \sin \delta$ and thereafter moves in a keplerian ellipse of eccentricity

$$e \sim \frac{U \sin \delta}{nr} = \frac{4}{3} \frac{m}{M} \left(\frac{r}{x} \right)^2 \quad (7)$$

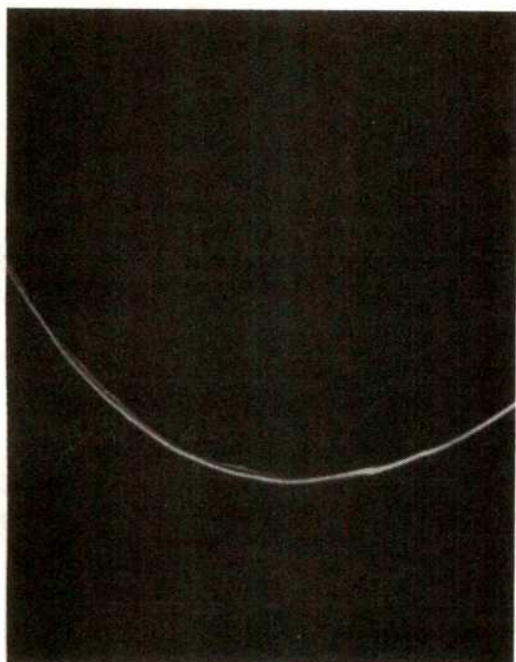


Fig. 1 This image, taken at a range of 750,000 km by Voyager 1, shows the highly complex structure of Saturn's F-ring. Two narrow 'braided' bright rings that trace distinct orbits are evident. Visible is a broader, very diffuse component, about 35 km wide. Also seen are 'knots' which are probably local clumps of ring material, but may be small satellites. Image courtesy JPL/NASA.

(see Fig. 2). A more accurate calculation by Julian and Toomre⁵ shows that the coefficient in this equation is 2.24 and not 4/3.

In a frame co-rotating with the perturbing satellite, all particles initially moving in circular orbits must follow identical paths after encounter. Thus each satellite generates a standing wave of amplitude

$$A = er \quad (8)$$

and wavelength

$$1 = \frac{2\pi U}{n} = 3\pi x \quad (9)$$

(see Fig. 2). In the inertial frame, each particle moves in an independent keplerian ellipse, but the pericentres of these elliptical orbits and the phases of the particles on the orbits are such that the locus of the particles is a sinusoidal wave which moves through the ring with the angular velocity of the perturbing satellite.

Because the tidal torques on the ring are equal, it follows from equations (6) and (8) that the amplitudes of the two waves are equal. For the F-ring of Saturn

$$A \sim 54 \left(\frac{m/M}{9 \times 10^{-9}} \right) \left(\frac{1,000 \text{ km}}{x} \right)^2 \text{ km} \quad (10)$$

and may be comparable with the width of the ring ($W \leq 30 \text{ km}$) (ref. 3). The interaction of these two waves, which probably have quite different wavelengths, could produce novel phenomena such as beats.

Resonance

Because $W \ll x$, the perturbations of the ring particle orbits will be highly coherent and although particle collisions will act to damp the waves, the process may be slow (and for this reason the magnitude of the torque given by equation (3) may be an overestimate). If the waves can survive from one encounter to the next, then we must consider the possibility of resonance.

This will occur if the ring circumference is an integral number of wavelengths, that is, if

$$\frac{2r}{3x} = p + 1 \quad (11)$$

where p is an integer. Consecutive perturbations will then be in phase and a wave of amplitude significantly greater than A may result.

The exact resonance condition can also be written

$$\frac{n'}{n} = \frac{p}{p+q} \quad \text{or} \quad \frac{p+q}{p} \quad (12)$$

where n' is the mean motion of the perturbing satellite and $q = 1$ for the type of resonance (first-order) that we are considering. As p can take any integral value ≥ 1 , a large number of resonances may occur. These are separated by a radial distance s , where

$$s = \frac{3}{2} \frac{x^2}{r} \quad (13)$$

For the satellites which bound the F-ring,

$$x_o \sim 2,500 \text{ km} \quad (14)$$

$$x_i \sim 800 \text{ km} \quad (15)$$

(ref. 2). Hence

$$s_o \sim 68 \left(\frac{x_o}{2,500 \text{ km}} \right)^2 \text{ km} \quad (16)$$

and

$$s_i \sim 7 \left(\frac{x_i}{800 \text{ km}} \right)^2 \text{ km} \quad (17)$$

These two sets of resonances will have quite different strengths. The amplitudes of the waves generated by a single encounter will be equal, but the larger satellite, which must be more distant

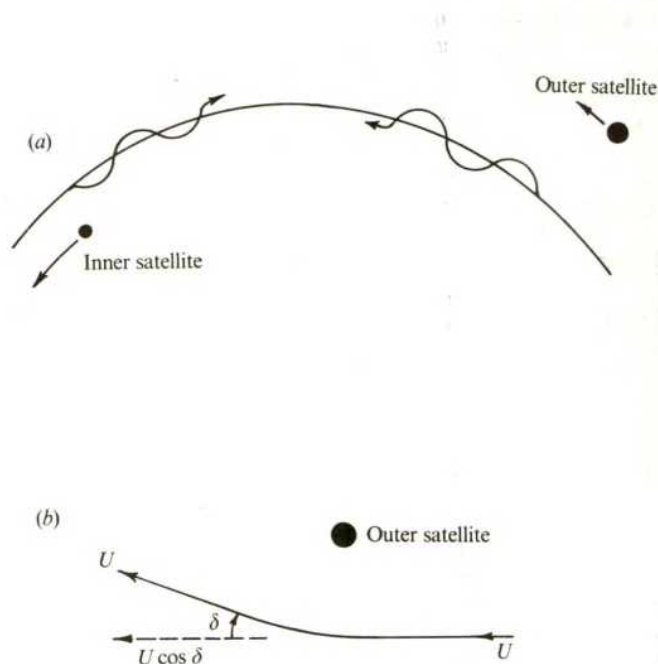


Fig. 2 Confinement of a narrow ring by two small satellites¹. *a*, Angular velocity decreases with increasing distance from the planet. Thus ring particles leading the outer satellite and lagging the inner satellite will be forced to move in eccentric orbits. Waves are excited which co-rotate with the perturbing satellite. *b*, The amplitude of the waves can be estimated from the scattering angle δ (ref. 2).

from the ring than the smaller satellite (see equation (6)), will have the larger relative velocity U and the time between encounters will be correspondingly shorter. Thus one would expect the resonances associated with the larger, outer satellite to dominate the structure of the ring. But these resonances are separated by a distance which is greater than the width of the ring. Thus, this satellite could have one, but only one, first-order resonance within the ring. The other possibility is that the width of the ring is < 7 km, and that a single first-order resonance associated with the smaller satellite exists within the ring.

For particles close to the exact resonance, the resonance condition is

$$pn - (p+1)n' + \dot{\omega} = 0 \quad (18)$$

where $\dot{\omega}$ is the pericentre of a ring particle orbit. If the particle is locked in resonance, and the forced eccentricity is small, then the pericentres of the ring particle orbits are not aligned, rather $\dot{\omega}$, $\dot{\omega}'$ and the mean longitude, or phase, are such that conjunctions of a particle and the satellite always occur at an apse of that particle's orbit⁶. The magnitude of the forced eccentricity, in the absence of damping, is given by⁶

$$e = \left| \frac{1/2(m/M)f(p)n'}{(p+1)n' - pn} \right| \quad (19)$$

Thus e increases markedly as the exact resonance is approached. At the exact resonance, the phase of the response changes by 180° . Similar behaviour is observed in any driven harmonic oscillator. For particles outside the exact resonance conjunction always occurs at apocentre, whereas for particles on the inside of the exact resonance conjunction always occurs at pericentre. Thus, the satellite excites a wave pattern of equally spaced loops which co-rotate with the perturbing satellite (see Fig. 3).

If p is large, then the inside width W_1 of a loop, that is, the maximum width of the gap generated in the ring, is given by⁷

$$W_1 = 2.96 \left(\frac{m}{M} \right)^{1/2} r \quad (20)$$

Thus, the effect of a given first-order resonance is determined solely by the mass of the perturbing satellite and is independent of both p and x . For the F-ring of Saturn,

$$W_1 = 39 \left(\frac{m/M}{9 \times 10^{-9}} \right)^{1/2} \text{ km} \quad (21)$$

The magnitude of this gap should be compared with the amplitude A of the short-period perturbations (see equation (10)). Two regimes can be clearly identified. If $A \gg 1/2 W_1$, then loop formation will be forever frustrated by the large perturbations suffered by the ring particles at each passage of the perturbing satellites. However, if $1/2 W_1 \gg A$, then loop formation is inevitable.

From equations (10) and (20), it follows that for $1/2 W_1 > A$, we must have

$$p < 56 \left(\frac{9 \times 10^{-9}}{m/M} \right)^{1/4} \quad (22)$$

or

$$x > 1,700 \left(\frac{m/M}{9 \times 10^{-9}} \right)^{1/4} \text{ km} \quad (23)$$

Thus, provided that the perturbing satellite is not so distant that first-order resonance is not possible, that is, provided that

$$\frac{n'}{n} > \frac{1}{2}, \quad (n' < n) \quad (24)$$

then the more distant the perturbing satellite the more likely is loop formation. However, if the unperturbed width of the ring is W , then the probability P that a first-order resonance will exist

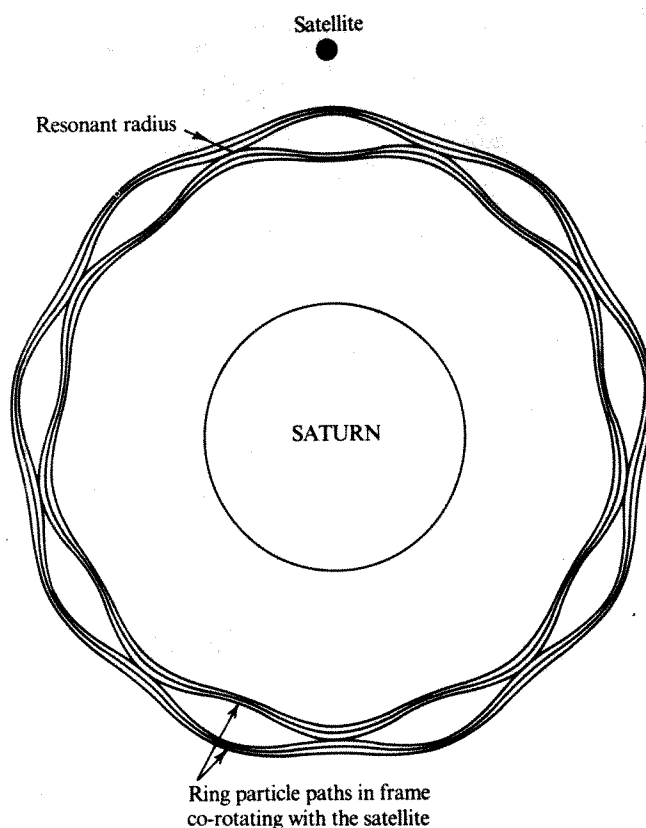


Fig. 3 Ring particle paths in a frame co-rotating with the perturbing satellite. Resonant gravitational interactions at radii in the ring where the ratio of the satellite and ring particle mean motions are close to $p/(p+1)$, where p is an integer, generate a system of waves which are stationary in the rotating frame. The waves on opposite sides of the exact resonance have a phase difference of 180° , hence the resultant wave pattern consists of $p+1$ equally spaced loops. This could account for the braided appearance of the F-ring of Saturn.

within the ring is given by

$$P = \frac{W}{s} = \frac{2rW}{3x^2} \quad (25)$$

and for the F-ring of Saturn,

$$P = 0.96 \left(\frac{W}{30 \text{ km}} \right) \left(\frac{1,700 \text{ km}}{x} \right)^2 \quad (26)$$

This suggests that loop formation is both possible and likely for satellites more distant than 1,700 km but not more distant than, say 5,000 km. According to early Voyager 1 reports², S13 orbits 2,500 km outside the F-ring, thus I consider that it is this satellite which is responsible for the 'braiding' of the F-ring.

Resonance is necessary to produce the bifurcation of particle orbits which is the essence of loop formation. The resonance need not be in the middle of the ring, in which case the two 'braids' will not necessarily be of equal brightness. Also, the short-period perturbations will not be negligible and, if the satellite orbits are not perfectly coplanar with the mean ring plane, waves both in the plane and out of the plane will be continuously excited by these satellites. The passage of these waves through the loops will give rise to shock phenomena associated with close-packing of ring particles⁷ and, possibly, because the ring is probably just outside Roche's limit for satellite formation, the formation of temporary particle clumps. Furthermore, because of these waves, the length and shape of

the loops will continuously change, and the structure of the ring, although periodic, will be highly complex and variable.

This theory cannot account for the third, diffuse ring. Perhaps this ring consists of small (micrometre-sized) particles which are generated in the main ring by erosion due to sputtering and meteoroid impacts and then removed by electromagnetic forces.

This article was inspired by comments from P. Goldreich and S. Tremaine during the Voyage 1 Saturn encounter. I thank C. D. Murray, R. Farouki and T. Gold for useful contributions. The data are from early Voyager Bulletin Mission Status Reports and I thank the Voyager Imaging Team (leader B. A. Smith) for

their generosity in releasing this information before their own publication. This research was supported by NSF grant AST 79164-74-A01.

Received 16 December 1980; accepted 18 February 1981.

1. Goldreich, P. & Tremaine, S. *Nature* **277**, 97-99 (1979).
2. *Voyager Bull., Mission Status Rep.* No. 56 (31 October 1980).
3. Smith, B. A. *et al. Science* (in the press).
4. Lin, D. N. C. & Papaloizou, J. *Mon. Not. R. astr. Soc.* **186**, 799-812 (1979).
5. Julian, W. H. & Toomre, A. *Astrophys. J.* **146**, 810-830 (1966).
6. Greenberg, R. *Astr. J.* **78**, 338-346 (1973).
7. Dermott, S. F., Farouki, R. & Murray, C. D. (in preparation).

Sequence and organization of the human mitochondrial genome

S. Anderson, A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin*, I. C. Eperon, D. P. Nierlich*, B. A. Roe*, F. Sanger, P. H. Schreier*, A. J. H. Smith, R. Staden & I. G. Young*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The complete sequence of the 16,569-base pair human mitochondrial genome is presented. The genes for the 12S and 16S rRNAs, 22 tRNAs, cytochrome c oxidase subunits I, II and III, ATPase subunit 6, cytochrome b and eight other predicted protein coding genes have been located. The sequence shows extreme economy in that the genes have none or only a few noncoding bases between them, and in many cases the termination codons are not coded in the DNA but are created post-transcriptionally by polyadenylation of the mRNAs.

MITOCHONDRIA have a separate autonomously replicating DNA genome (for a review see ref. 1). Although most of the proteins present in the mitochondrion are encoded by nuclear DNA, a few are coded for by mitochondrial (mt) DNA and are synthesized by the separate mt translation system. The rRNAs and tRNAs of this system are mitochondrially encoded and it is likely that the other components, although nuclear coded, are distinct from those in the cytoplasmic translation system. The enzymes specific for the unique mitochondrial transcription and replication processes are probably also nuclear coded. We have recently shown that mammalian mitochondria have a different genetic code^{2,3} as have other mitochondria systems that have been studied⁴⁻⁷, and a different decoding mechanism⁷⁻⁹.

Replication of mammalian mtDNA proceeds by initiation of heavy (H) strand synthesis at a specific origin^{10,11} resulting in the formation of a displacement (D) loop with a newly synthesized H strand of ~680 bases known as 7S DNA. In mouse cells a large proportion of mtDNA exists in the D-loop form but only a few of these nascent H strands grow to full length. Initiation of light (L) strand synthesis is at a specific origin¹² and does not occur until this region has been exposed by H-strand synthesis.

Transcription of mammalian mtDNA is also unique in that both strands seem to be completely transcribed from promoters situated in the D-loop region^{13,14}. These primary transcripts are then processed to give the 12S and 16S rRNAs, tRNAs and a number of presumptive mRNAs, which are not capped but are polyadenylated¹⁵. The small size of the ribosomal RNAs, the chloramphenicol sensitivity of mitochondrial protein synthesis and the use of formyl-methionine in initiation of translation

have led to the suggestion that the mitochondrial genetic system arose by endosymbiosis of primitive bacteria within a host cell. The sequence of the rRNA genes, however, shows only low homology with present-day prokaryotic rRNAs¹⁶.

To understand this genetic system more fully we have determined the complete DNA sequence of both the human and the bovine mitochondrial genomes. We report here the sequence of human mtDNA and an interpretation of this sequence in terms of its coding capacity, gene arrangement and gene expression. The bovine sequence will be published elsewhere.

In the accompanying articles Ojala *et al.*¹⁷ and Montoya *et al.*¹⁸ report the sequence analysis of the stable polyadenylated RNAs. The comparison of the DNA and RNA sequences reveals the unique nature of the precise transcript processing in mammalian mitochondria.

Organization of the genome

The DNA sequence is shown in Fig. 1 along with the predicted coding sequences. This is presented schematically in Fig. 2. The identification of the genes and their precise location in the DNA sequence are based on several unique features, the evidence for which is discussed below and in the accompanying articles^{17,18}. The sequence shows extreme economy of organization in that there are none or very few noncoding bases between adjacent genes except for the D-loop region where we have been unable to assign any coding function. The sequence shown is that of the L strand which we define as the main coding strand containing the sense sequence of the rRNAs and most of the tRNAs and mRNAs. Thus, these RNAs are transcribed from (and therefore hybridize to) the H strand.

Origins of DNA replication

The 5' end of the major 7S DNA species found in the D-loop of HeLa cell mtDNA¹⁰ corresponds to nucleotide 191 of the DNA sequence. Other human 7S DNA species having different 5' ends have also been reported^{11,19,20}, but the precise location of these termini in the sequence presented here is not known. The size of the 7S DNA (≤ 680 nucleotides) and mapping data^{10,19,20}

*Present addresses: Department of Biochemistry and Biophysics, School of Medicine, University of California, San Francisco, California 94143, USA (J.D.); Department of Microbiology, College of Letters and Science, University of California, Los Angeles, California 90024, USA (D.P.N.); Chemistry Department, Kent State University, Kent, Ohio 44242, USA (B.A.R.); Institut für Genetik der Universität, 5 Köln 41, Weyertal 121, FRG (P.H.S.); Department of Biochemistry, John Curtin School of Medical Research, Australian National University, PO Box 334, Canberra City, ACT 2601, Australia (I.G.Y.).

Table 1 Human mitochondrial genetic code

Phe	UUU	77	Ser	UCU	32	Tyr	UAU	46	Cys	UGU	5
	UUC	141		UCC	99		UAC	89		UGC	17
Leu	UUA	73		UCA	83	Ter	UAA	—	Trp	UGA	93
	UUG	16		UCG	7		UAG	—		UGG	11
Leu	CUU	65	Pro	CCU	41	His	CAU	18	Arg	CGU	7
	CUC	167		CCC	119		CAC	79		CGC	25
	CUA	276		CCA	52	Gln	CAA	81		CGA	29
	CUG	45		CCG	7		CAG	9		CGG	2
Ile	AUU	125	Thr	ACU	51	Asn	AAU	33	Ser	AGU	14
	AUC	196		ACC	155		AAC	131		AGC	39
Met	AUA	167		ACA	133	Lys	AAA	85	Ter	AGA	—
	AUG	40		ACG	10		AAG	10		AGG	—
Val	GUU	30	Ala	GCU	43	Asp	GAU	15	Gly	GGU	24
	GUC	49		GCC	124		GAC	51		GGC	88
	GUA	70		GCA	80	Glu	GAA	64		GGA	67
	GUG	18		GCG	8		GAG	24		GGG	34

The genetic code of human mtDNA differs from the universal code in that UGA codes for tryptophan and not termination, AUA codes for methionine not isoleucine, and AGA and AGG are termination rather than arginine codons. In addition, AUA and possibly AUU are initiation codons as well as AUG. Boxes of four codons are each read by a single tRNA with U in the first position of the anticodon; boxes of two codons are read by tRNAs with G:U wobble anticodons. The number of methionyl tRNAs and their codon response are unclear and discussed in the text. Also shown are the total number of each amino acid codon found in the genes and predicted genes shown in Figs 1 and 2. Codons ending in A and C predominate (36.37 and 40.93%, respectively) and those ending in U and G are used less frequently (16.33 and 6.36%, respectively).

indicate that the 7S DNA, and presumably the D-loop, do not extend into the tRNA^{Pro} gene. Although H-strand synthesis during replication initiates in the D-loop region, data from mouse cell mitochondria indicate that most of the 7S DNA molecules do not participate in DNA replication but rather are synthesized and then rapidly lost from the D-loop²¹.

The ~1,100-base pair region of the human mitochondrial genome between the tRNA^{Pro} and tRNA^{Phe} genes contains no large open reading frames in either strand. Furthermore, except for a small 7S RNA¹⁵, no large stable polyadenylated transcript has been found to map in this region. Unless RNA splicing is invoked it is extremely unlikely that this part of the genome codes for a major protein species. In the region covered by the D-loop, several blocks of nucleotide sequence homology exist between the human, bovine (S.A. *et al.*, unpublished results) and rat mitochondrial genomes²². However, the DNA separating these blocks is rather variable in length and exhibits no significant homology. The region between the 5' end of the 7S DNA and the tRNA^{Phe} gene seems to be one of the least conserved regions of the entire genome. This was initially detected by electron microscopy of heteroduplexes of sheep and goat mtDNAs²³ and has been confirmed by the DNA sequence comparisons. Very little sequence homology between the human, rat²² or bovine mitochondrial genomes can be detected in this region. The size of the region also seems to vary much

more than do intergenic spaces in other parts of the mtDNA, being 385 nucleotides in the human placenta, 291 nucleotides in rat liver and (based on alignment with the human and rat sequence) ~180 nucleotides in beef heart. This region may contain control signals for transcription of the mitochondrial genome.

L-strand synthesis during mtDNA replication initiates two-thirds of the way around the genome from the region of H-strand synthesis²⁴. In mouse cell mitochondria the DNA sequence immediately surrounding the L-strand origin is a simple inverted repeat that may form a hairpin loop structure when the L-strand origin is exposed by H-strand displacement synthesis during DNA replication¹². An analogous and highly homologous presumptive L-strand origin exists in the human mtDNA sequence (nucleotides 5,730–5,763) within a cluster of five tRNA genes. Interestingly, the differences between the human and mouse inverted repeats occur symmetrically, so that if secondary structure did exist, full base-pairing in the stem region would be maintained.

Ribosomal RNA genes

Eperon *et al.*¹⁶ were able to align the DNA sequence with the sequences of the 5' ends^{15,25} of the 12S and 16S rRNAs and to show that there were no noncoding bases between the tRNA^{Phe} gene and the 12S rRNA gene, and between the tRNA^{Val} gene

Fig. 1 The DNA sequence of human mitochondria (pages 459–461). The sequence of the L strand is numbered arbitrarily from the *Mbo*I-5/7 boundary⁵⁶ in the D-loop region. The sequence is displayed using the program TRANMT of Staden (unpublished) and shows the position of the genes for the tRNAs (boxed) and the rRNAs (underlined). The unidentified reading frames (URF 1–6) and identified protein coding genes are translated into the one-letter amino acid code (termination codons are denoted by an asterisk). The genes for cytochrome *c* oxidase subunits I–III are abbreviated COI–III and *cyt b* denotes the gene for cytochrome *b*. Which strand contains the sense sequence of the genes is denoted by → (L strand) and ← (H strand). The ends of the presumptive mRNAs transcribed from the H strand as determined by RNA sequence analysis^{17,18} are also shown below the sequence. The 5' end of the nascent H-strand ('7S') DNA found in the D-loop is at nucleotide 191 (ref. 10). The origin of L-strand synthesis during DNA replication¹² is indicated by overlining of the possible base-pairing regions. The DNA sequence was derived mainly from a single human placenta mtDNA preparation⁵⁶ but some regions were determined on HeLa cell mtDNA. Restriction enzyme fragments of placental mtDNA were cloned in pBR322⁵⁶. Single-stranded template DNA was prepared by the exonuclease III procedure⁵⁷ and by cloning in bacteriophage M13 (refs 58–61). The DNA sequence was determined by the chain termination method of Sanger *et al.*^{62,63}. The sequence was analysed using the computer programs of Staden (refs 31, 55, 64 and unpublished). Nucleotides 10, 934–5, 14,272 and 14,365 were ambiguous and their identity assumed to be the same as in the bovine mtDNA sequence (S.A., M.H.L.B., A.R.C., I.C.E., F.S., I.G.Y., unpublished). The base composition of the L strand is 24.7% T, 31.2% C, 30.9% A and 13.1% G.

Figure 1

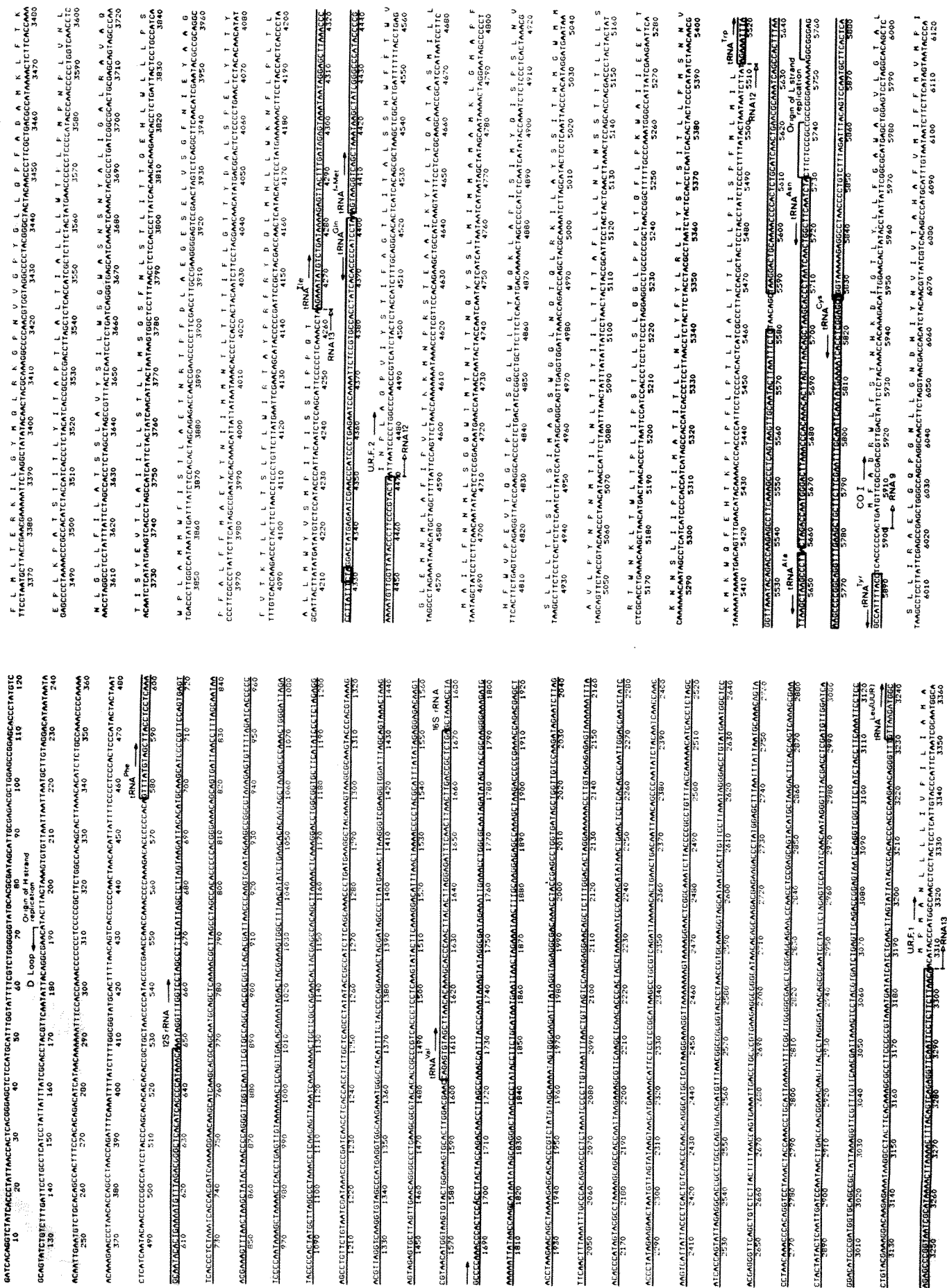
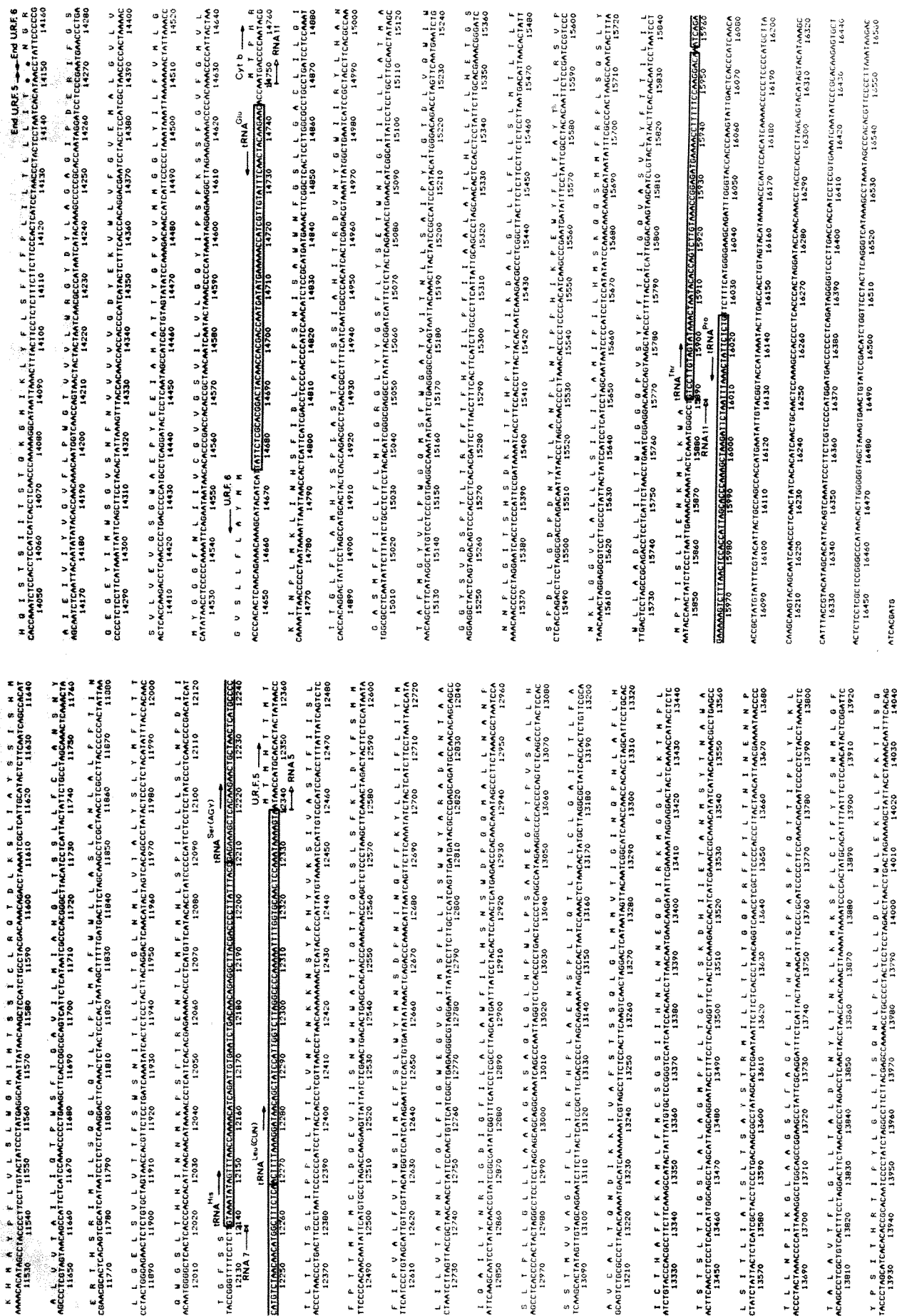


Figure 1



Figure 1



and the 16S rRNA gene. However, the 3' ends of the rRNA genes could only be predicted to be very close to the 5' ends of the flanking tRNA^{Val} and tRNA^{Leu} genes. It has since been shown that the mouse small subunit rRNA gene abuts the 3' tRNA^{Val} gene^{26,27} and alignment of a hamster large subunit rRNA sequence with the murine gene sequence suggests that there are also no noncoding bases between the large subunit rRNA gene and the 3' tRNA^{Leu} gene²⁸. Thus, the 12S rRNA and 16S rRNA and the flanking tRNAs in a primary transcript could be released by cleavage precisely at the 5' and 3' ends of the tRNAs without the need for further processing.

The high concentration of nascent transcripts from this region of the genome has led to the suggestion that there is a transcription attenuator in this region²⁰. This is supported by the observation that the hamster large subunit rRNA ends and is polyadenylated at any of four nucleotides within the last five nucleotides of the gene²⁸, which contrasts with the precise termini of the 12S rRNA and mRNAs^{29,30} which are presumed to be generated by processing of a primary transcript. It is possible that the 12S rRNA is also adenylated at the 3' terminus after transcription and processing: a single addition has been proposed for the hamster small subunit rRNA by comparison with the murine gene sequence³⁰.

The sequences of the rRNA genes were interpreted to provide support for a model in which mt ribosomes did not have sequence-dependent signals for recognition of the correct initiation codon but rather initiated at the first initiation codon in the mRNA¹⁶. This is consistent with the known 5'-end sequences of the mRNAs¹⁸ although not with the observation of single transcripts covering the URF A6L/ATPase 6 and URF 4L/URF 4 reading frames (Fig. 2) if the expected proteins are expressed.

The tRNA genes

Twenty-two tRNA genes have been identified in the human mtDNA sequence either by visual inspection or using the program TRNA³¹. Equivalent genes have been located in the bovine mtDNA sequence and most of these have been confirmed by direct sequencing of bovine mitochondrial tRNAs (B. A. Roe, E. Y. Chen, P. W. Armstrong and J. F. H. Wong, unpublished). We have been unable to substantiate a previous report⁸ which tentatively identified a gene for tRNA^{Met}. Apart from the tRNA^{Met}, these 22 tRNAs are sufficient to read all codons using a mechanism unique to mitochondrial systems⁷⁻⁹. No mammalian mt tRNAs have been shown to be imported from the cytoplasm (ref. 32 and B. A. Roe and E. Y. Chen, unpublished results). It had previously been assumed that the minimum number of tRNAs required to read all codons would be 32 for the 'universal' genetic code using G:U or I:A/C/U wobble³³. In human mitochondria the tRNAs for the two-codon families, that is, those genetic code boxes with two codons for one amino acid, have G:U wobble anticodons. However, only one tRNA is found for each of the four-codon families⁸ (see Table 1). These tRNAs have a U in the first position of the anticodon although each must read all four codons in its respective family. This is probably accomplished by U:N wobble⁸, although a two-out-of-three mechanism might operate³⁴. Although the modification pattern of the U in the first position of the anticodon of mammalian mt tRNAs is now known, it is likely that a similar situation exists to that found in *Neurospora* mitochondria⁷, whereby the tRNAs for the four-codon families have an unmodified U, but when the U is required for reading the two-codon families it is modified. This modification presumably prevents such tRNAs from reading codons ending in U or C.

Two species of mammalian mt methionyl tRNAs have been identified, a tRNA^{Met} and a tRNA^{Met} (refs 32, 35). Hybridization of these tRNAs to mtDNA showed partial additivity, indicating the existence of two separate genes³². In the DNA sequences we have only been able to identify one tRNA^{Met} gene, presumably coding for the initiator tRNA, based on its anticodon sequence CAT and the close homology of its anticodon arm with other initiator tRNAs (see ref. 36). The tRNA anti-

codon sequence CAU should normally be specific for AUG but as discussed below we believe that AUA and possibly AUU are also initiation codons in mammalian mitochondria. To recognize AUA (and AUU?) as well as AUG, the C in the first position of the anticodon would probably have to be modified. The AUA-specific tRNA^{Met} of *Escherichia coli* and of bacteriophage T4 have modified C residues in the first position of the anticodon^{37,38}, and hence represent examples of specific recognition of AUA by a tRNA with a CAU anticodon. Although potential tRNA^{Met}-like structures are found in the human and bovine sequences, none of these is particularly convincing nor are they conserved in the corresponding sequence. Also, extensive sequence analysis of the bovine mt tRNAs has not revealed any candidate for a tRNA^{Met} (B. A. Roe, E. Y. Chen, P. W. Armstrong and J. F. H. Wong, personal communication). If the gene exists in the mt genome it may have an unusual structure that is preventing its detection, for example, like that of tRNA^{Ser}_{AGU/C} which completely lacks the D arm³⁹. The other possibility is that one gene can give rise to both a tRNA^{Met} and a tRNA^{Met} by differential modification, although the hybridization data of Aujame and Freeman³² suggest otherwise. We are at present trying to clarify this problem by hybridization of labelled methionyl tRNA to restriction fragments of mtDNA.

The sequences of the mammalian mt tRNAs are unusual in that all, except tRNA^{Leu}_{UUR}, lack some or all of the following features found in other tRNAs: (1) the universal sequence G-T-ψ-C-R-A, (2) the constant seven-base length of the 'TψC' loop which in mammalian mt tRNAs varies between three and nine bases, and (3) the constant bases A₁₄, G₁₅ and G₁₈G₁₉ and the connections with U₈ and U₄₈ (yeast tRNA^{Phe} numbering system, see ref. 40). Thus, the mt tRNAs are apparently stabilized by fewer tertiary interactions. The most extreme case is that of tRNA^{Ser}_{AGU/C}, which lacks the D arm³⁹. Also, considering the similarity of homologous cytoplasmic tRNA species from different animals it is surprising that substantial variations are found when specific tRNAs from human and bovine mitochondria are compared. A compilation and comparison of these tRNAs will be published elsewhere. These missing or different tertiary interactions may mean that the mammalian mitochondrial tRNAs have more freedom than cytoplasmic tRNAs to evolve in certain areas, perhaps reflecting a different interaction with the ribosome. The role of the tRNAs in processing of the primary transcript is discussed below.

Table 2 Characteristics of human mt genes and comparison with bovine mt genes

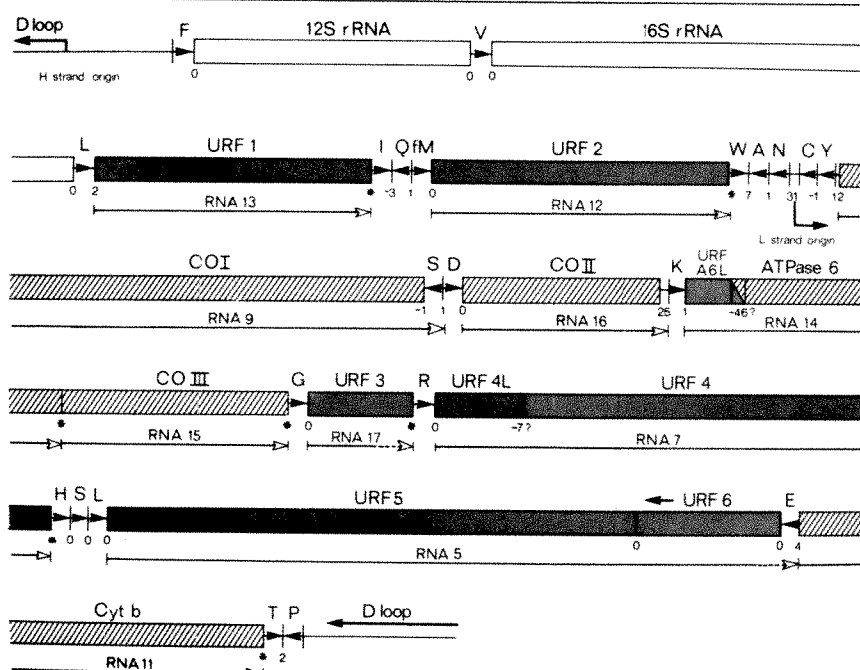
Gene	Protein length	Molecular weight	% Amino acid conservation	Initiation and termination codons	
				Human	Bovine
URF1	318	35,600	75.79	ATA *	ATG *
URF2	347	38,900	62.82	ATT *	ATA *
COI	513	57,000	91.23	ATG AGA	ATG TAA
COII	227	25,500	72.69	ATG TAG	ATG TAA
URF A6L	68	7,900	51.52	ATG TAG?	ATG TAA?
ATPase 6	226	24,800	77.88	ATG *	ATG *
COIII	261	30,000	86.97	ATG *	ATG *
URF3	115	13,200	73.91	ATA *	ATA *
URF 4L	98	10,700	73.47	ATG TAA?	ATG TAA?
URF4	459	51,400	74.07	ATG *	ATG *
URF5	603	66,600	69.49	ATA TAA	ATA TAA
URF6	174	18,600	62.64	ATG AGG	ATG TAA
Cytochrome b	380	42,700	78.10	ATG *	ATG AGA

The predicted length and molecular weight of the expected products of the human mt genes and unidentified reading frames were calculated from the DNA sequence using the computer programs of Staden (ref. 55 and unpublished results). Also shown is the percentage identical amino acids conserved between the human and bovine sequences calculated using the program AACHAN (R. Staden, unpublished). The predicted initiation and termination codons of the human and bovine genes are shown; * indicates that the termination codon is created post-transcriptionally by polyadenylation (see text).

Protein coding genes

Thirteen significant reading frames have been identified in the human mtDNA sequence, of which five have been assigned to

Fig. 2 Organization of the genome. A schematic representation of the sequence shown in Fig. 1. The tRNA genes are identified by the one-letter amino acid code and their sense sequence by either $\rightarrow$ (L strand) or $\leftarrow$ (H strand). Unidentified reading frames 1–5 (URF1–5) and the identified protein genes for cytochrome *c* oxidase subunits I–III (COI–III), ATPase 6 and cytochrome *b* (cyt *b*) all have their sense sequence on the L strand; URF6, however, is H-strand coded. The number of noncoding bases between genes is given below the junctions of the genes. A minus number denotes an overlap and an asterisk indicates that there are no noncoding bases at that point and that the termination codon for the preceding gene is created in the mRNA by polyadenylation (see text). The numbered RNAs are the presumptive mRNAs transcribed from the H strand which were aligned with the DNA sequence by hybridization to restriction fragments and by sequence analysis of their 5' and 3' ends (see accompanying articles^{17,18} for a more detailed transcription map showing the positions of the L-strand transcripts and the precursors to the H-strand transcripts).



known mitochondrially synthesized proteins. Only reading frames of at least 40 amino acids conserved with at least 50% amino acid homology in the bovine sequence were considered. The gene for cytochrome *c* oxidase subunit I (COI) was identified from an N-terminal amino acid sequence of the bovine protein (J. E. Walker, personal communication) and confirmed by comparison with the yeast mt COI gene. COII was identified from the complete amino acid sequence of the bovine protein⁴¹ and COIII by homology with the yeast gene⁴² and by an N-terminal sequence of the bovine protein (G. Buse, personal communication). The genes for ATPase 6 and cytochrome *b* were identified by homology with the corresponding yeast genes^{43,44} and also by amino acid sequences of bovine cytochrome *b* (W. Machleidt, personal communication). We have found no homology between the bovine amino acid sequence of ATPase 9 (E. Wachter, personal communication) and any of the reading frames, so that, as in *Neurospora*⁴⁵, this protein must be nuclear coded; in yeast it is mitochondrially coded^{46,47}.

In regions representing identified genes the respective human and bovine sequences are highly homologous. As shown in Table 2, there is ~80% amino acid conservation, which is 5% higher than the nucleotide conservation, reflecting the degeneracy of the genetic code. Over half the codons for the conserved amino acids are different in the third (degenerate) position or, in the case of the leucine codons UUR and CUN, in the first and third positions.

We have found eight unidentified reading frames (URFs) that are conserved in human and bovine mtDNA. On average, these show ~70% nucleotide and amino acid conservation (Table 2), again with approximately half of the conserved amino acids having changed codons. This type of homology is unlikely to have occurred by chance: for example, if the reading frames are compared out of phase in the +1 and +2 frames, the amino acid homology, including all the termination codons, is 35.4% and the percentage of conserved amino acids with changed codons drops to 4.5%. Therefore, we believe that most, if not all, of these reading frames represent genes for mammalian mitochondrial proteins. Further evidence in support of this is provided by the mapping and sequence analysis of the polyadenylated presumptive mRNAs reported in the accompanying articles^{17,18}. These RNAs correspond exactly to the reading frames of URF1, URF2, COI, COII, URF A6L and ATPase 6, COIII, URF3, URF 4L and URF4, URF5 and the antisense of URF6, and cytochrome *b* as shown in Figs 1 and 2. In addition, there are large L-strand transcripts covering the H-strand-coded URF6, and the number and sizes of these reading frames are in fairly good agreement with the band pattern obtained by

gel electrophoresis of HeLa cell mt proteins produced in the presence of emetine, an inhibitor of cytoplasmic protein synthesis (J. E. Walker, personal communication).

Two of the H-strand transcripts contain more than one reading frame, RNA14 (URF A6L and ATPase 6) and RNA7 (URF 4L and URF4). These present an anomaly and it is not known whether the reading frames are separated by a further processing event of these RNAs. The same organization is found in bovine mtDNA, making it unlikely that the breaks in the reading frames are due to sequence errors. A possible explanation is that these frames contain short intervening sequences which, when excised, leave only a single reading frame. However, using nuclease S₁ protein experiments, Ojala *et al.*⁴⁸ found no evidence for intervening sequences in the mtDNA regions coding for the polyadenylated RNAs.

Initiation codons

Although all the identified genes (except ATPase 6) and some of the URFs have an AUG initiation codon within the first six bases of their mRNA start, others do not. The first ATG in bovine URF2 is 320 codons downstream from the RNA start in a reading frame of 347 codons. Also, in human and bovine URF3 the first AUG is, respectively, 89 and 87 codons downstream in a reading frame of 115 codons. Ignoring the cases of ATPase 6 and URF4, which are the second reading frames in their presumptive mRNAs, and also the case of human URF2, which is discussed separately below, all reading frames that do not have an AUG near the 5' end of the RNAs have an AUA in an equivalent position. Thus, if these reading frames are translated completely, as their homology suggests they are, AUA must function as an initiation codon as well as AUG. AUA has already been shown to code for internal methionine as well as AUG^{2,3}. In the case of human URF2 there is an AUU in the same position as the AUA in bovine URF2. Thus, to translate completely this reading frame we must argue that, like AUA, AUU can also be an initiation codon in mitochondrial genes.

Termination of translation

Over half of the presumptive mRNAs do not have UAA or UAG termination codons at the end of the reading frame (Table 2). In three cases, human COI and URF6 and bovine cytochrome *b*, termination probably occurs at AGA and AGG codons. These have been predicted to be termination codons and not arginine codons as in the universal genetic code⁸. This is based on the observation that only CGN arginine codons are used in all the reading frames, and AGA and AGG are only found at the ends of reading frames at the junction with the next

36. Wrede, P., Woo, N. H. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3289–3293 (1979).
37. Kuchino, Y., Watanabe, S., Harada, F. & Nishimura, S. *Biochemistry* **19**, 2085–2089 (1980).
38. Fukada, K. & Abelson, J. *J. molec. Biol.* **139**, 377–391 (1980).
39. de Bruijn, M. H. L. *et al. Nucleic Acids Res.* **8**, 5213–5222 (1980).
40. Grauss, D. H., Grüter, F. & Sprinzl, M. *Nucleic Acids Res.* **6**, r1–r19 (1979).
41. Steffens, G. J. & Busp, G. *Hoppe-Seyler's Z. physiol. Chem.* **360**, 613–619 (1979).
42. Thalenfeld, B. E. & Tzagoloff, A. *J. biol. Chem.* **255**, 6173–6180 (1980).
43. Macino, G. & Tzagoloff, A. *Cell* **20**, 507–517 (1980).
44. Nobrega, F. G. & Tzagoloff, A. *J. biol. Chem.* **255**, 9828–9837 (1980).
45. Sebald, W. & Wachter, E. in *29th Mossbacher Colloquium on Energy Conservation in Biological Membranes* (eds Schäfer, G. & Klingenberg, M.) 228–236 (Springer, Berlin, 1978).
46. Hengstens, L. A., Grivell, L. A., Borst, P. & Bos, J. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1663–1667 (1979).
47. Macino, G. & Tzagoloff, A. *J. biol. Chem.* **254**, 4617–4623 (1979).
48. Ojala, D., Merkel, C., Gelfand, R. & Attardi, G. *Cell* **22**, 393–403 (1980).
49. Barrell, B. G. *et al.* in *31st Mossbacher Colloquium* (eds Bücher, Th., Weiss, H. & Sebald, W.) (Springer, Berlin, in the press).
50. Robertson, M. D., Altman, S. & Smith, J. D. *J. biol. Chem.* **247**, 5243–5251 (1972).
51. Altman, S. & Smith, J. D. *Nature new Biol.* **233**, 35–39 (1971).
52. Attardi, G. *et al.* in *The Organisation and Expression of the Mitochondrial Genome* (eds Kroon, A. M. & Saccone, C.) (North-Holland, Amsterdam, in the press).
53. Battey, J. & Clayton, D. A. *J. biol. Chem.* **255**, 11599–11606 (1980).
54. Borst, P. & Grivell, L. A. *Cell* **15**, 705–723 (1978).
55. Staden, R. *Nucleic Acids Res.* **4**, 4037–4051 (1977); **5**, 1013–1015 (1978).
56. Drouin, J. *J. molec. Biol.* **140**, 15–34 (1980).
57. Smith, A. J. H. *Nucleic Acids Res.* **6**, 831–848 (1979).
58. Gronenborn, B. & Messing, J. *Nature* **272**, 375–377 (1978).
59. Schreier, P. H. & Cortese, R. *J. molec. Biol.* **129**, 169–172 (1979).
60. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).
61. Anderson, S., Gait, M. J., Mayol, L. & Young, I. G. *Nucleic Acids Res.* **8**, 1731–1743 (1980).
62. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
63. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107–110 (1978).

Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs

Julio Montoya, Deanna Ojala & Giuseppe Attardi

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

The 5'-end proximal sequences of all the putative mRNAs coded for by the heavy strand of HeLa cell mitochondrial DNA have been determined and aligned with the DNA sequence. All these mRNAs start directly at, or very near to, an AUG or AUA triplet, with the exception of one which starts at an AUU. The available evidence indicates that the terminal or subterminal AUGs and AUAs, and possibly also the terminal AUU, are initiator codons for the corresponding polypeptides. In most cases, the individual mRNA coding sequences are flanked on their 5' side by a tRNA gene, without any intervening nucleotide.

RECENT sequence analysis of human mitochondrial DNA (mtDNA) has revealed an extremely high degree of packing of genetic information in this DNA (ref. 1 and accompanying paper (ref. 26)). In agreement with these observations, a detailed transcription mapping study of HeLa cell mtDNA by the S_1 protection technique has shown that the sequences of the heavy (H) strand are almost completely saturated by the rRNAs, poly(A)-containing RNAs and tRNAs coded for by this strand^{2,3}. A particularly intriguing feature of the human mitochondrial gene organization is the frequent juxtaposition, or close proximity, of the individual protein-coding sequences, on their 5' side, with a tRNA gene or another reading frame. Thus, the gene for subunit II of cytochrome c oxidase (COII) has been shown to be immediately contiguous to a tRNA^{Asp} gene without any intervening nucleotide⁴. This unusual gene organization has raised the question of whether these putative mitochondrial mRNAs have, as all eukaryotic mRNA so far analysed^{5,6}, a 5' non-coding stretch, which may thus overlap the contiguous tRNA gene or reading frame, or whether they lack completely a leader sequence on their 5' side.

As a first approach to this question, the 5'-end proximal region of the putative COII mRNA has recently been sequenced and the sequence thus determined, aligned with the COII gene sequence⁷. The results clearly showed that the 5' end of the COII mRNA corresponds precisely with the first nucleotide of the COII coding sequence. These experiments left open the question of whether the complete lack of 5' non-coding nucleotides is a general feature of human mitochondrial mRNAs, reflecting stringent constraints in the initiation of mitochondrial protein synthesis or specific rules of RNA processing, or whether, on the contrary, the position of the initiator codon relative to the 5' end may vary in different mRNAs, possibly depending on the position of the adjacent tRNA gene or reading frame. We report here information bearing on these questions, obtained by sequencing a 5'-end proximal segment of all the H strand-coded, polysome-associated, poly(A)-containing RNAs, which are

presumably specific mRNAs, and by aligning the RNA sequences with the mtDNA sequence.

Isolation and 5'-end labelling of mitochondrial mRNAs

To isolate HeLa cell mitochondrial poly(A)-containing RNAs in adequate amounts for sequencing analysis, the micrococcal nuclease procedure for the purification of mitochondrial RNA⁸ was scaled up as previously described⁷. Figure 1a shows the electrophoretic pattern in a 1.4% agarose-CH₃HgOH slab gel, after ethidium bromide staining, of a large-scale preparation of mitochondrial oligo(dT)-bound RNA from micrococcal nuclease-treated organelles (from ~30 g of cells). One clearly recognizes the characteristic set of components previously

Table 1 5'-Terminal nucleotide analysis of *in vitro* labelled mitochondrial poly(A)-containing RNA species

Poly(A)-containing RNA species	% Of total ³² P radioactivity*			
	AMP	CMP	GMP	UMP
5	52.5	8.1	27.6	11.7
7	43.8	2.1	7.8	47.1
9	12.4	0.9	2.0	4.5
11	81.1	3.0	10.2	5.5
12	78.9	2.3	12.7	5.9
13	81.4	2.4	10.0	6.0
14	91.0	3.1	2.2	3.5
15	95.9	0.3	1.6	2.1
17	73.7	2.5	6.4	17.2

The 5'-terminal nucleotide was determined by exhaustive digestion of the *in vitro* labelled RNAs with nuclease P₁ (ref. 12, Calbiochem) and fractionation of the products on polyethyleneimine impregnated-cellulose TLC plates, as previously described⁸.

* The values refer to the radioactivity which migrated from the origin. The % of the input radioactivity recovered from the origin varied between 0.1 and 5%.

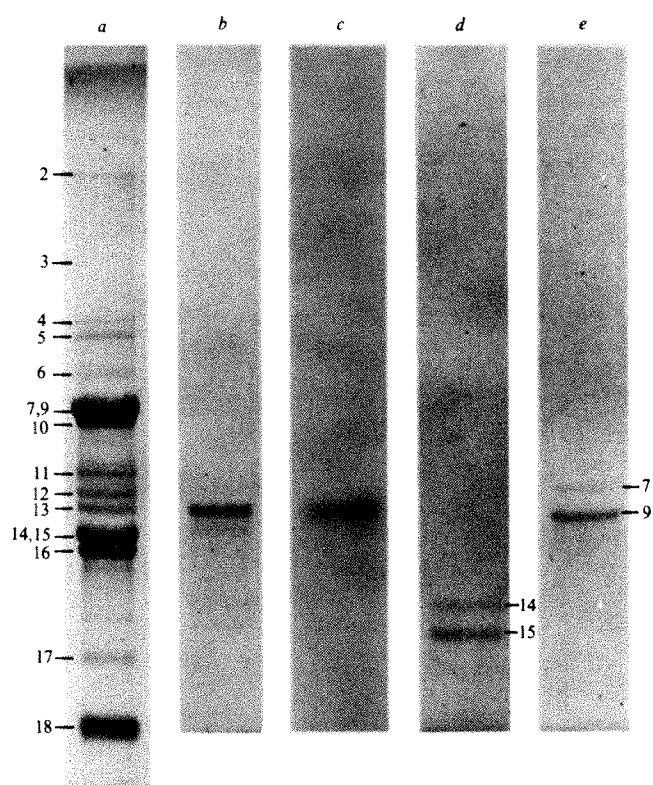


Fig. 1 Isolation and 5'-end labelling of mitochondrial poly(A)-containing RNA species. Lane *a*, electrophoretic pattern after ethidium bromide staining of the oligo(dT)-bound mitochondrial RNA fraction from 30 g of HeLa cells run through a 1.4% agarose-CH₃HgOH slab gel. Lanes *b* and *c*, RNA species 13 was eluted from several gel tracks similar to those shown in lane *a*, treated with RNase-free DNase and re-run on a gel in the same conditions (*b*), then eluted again, labelled at its 5' end with [γ -³²P]ATP and polynucleotide kinase after bacterial alkaline phosphatase treatment and re-run (*c*). Lanes *d* and *e*, re-run of RNA species 14 plus 15 (*d*) or 7 plus 9 (*e*) on 2% agarose-CH₃HgOH gels after elution from the gel track shown in lane *a* and treatment with DNase. Large-scale preparation of mitochondrial RNA by the micrococcal nuclease procedure^{7,8}, subsequent isolation and fractionation by 1.4% agarose-CH₃HgOH slab gel electrophoresis of the oligo(dT)-cellulose-bound RNA, elution, DNase treatment and re-run on agarose-CH₃HgOH slab gels of the individual components⁷, and 5' end labelling of the purified RNA species⁸ were as previously reported.

detected in ³²P-labelled preparations from HeLa cells⁹, with the exception of RNA species 1, which is extremely short-lived and present in mitochondria in very small amounts^{10,11}. Because of their relative abundance, their enrichment in partially purified polysomal structures⁹ and their relatively long half life^{10,11}, RNA species 5, 7, 9 and 11–17 in this set are probably specific mitochondrial mRNAs. Correlation of RNA sequence data, or of reading frames in mtDNA, with protein sequence data and with genetically characterized yeast gene sequences strongly supports the mRNA nature of several of these species^{1,2}. However, no direct *in vitro* demonstration of the messenger activity of these RNAs is available yet. Each of the putative mRNAs was eluted from the gel, treated with RNase-free DNase (to eliminate any contaminating DNA fragments deriving from the micrococcal nuclease treatment of the crude mitochondrial fraction) and re-run on a gel in the same conditions. As shown in Fig. 1*b* for RNA 13, a sharp band was observed at the expected position, and the RNA in this band was eluted. RNAs 7 and 9, and also RNAs 14 and 15, which were not separated from each other in the first electrophoretic run, could be resolved in a re-run on a 2% agarose-CH₃HgOH gel (Fig. 1*d*, *e*). Material purified as described above, from 40–250 g of cells, depending on the RNA species, was finally labelled at its 5' end with [γ -³²P]ATP and polynucleotide kinase after dephos-

phorylation with bacterial alkaline phosphatase, re-run on a CH₃HgOH-agarose gel (Fig. 1*c*) and eluted for the sequencing reactions.

RNA sequence analysis

The 5' termini of the isolated *in vitro* labelled RNA species were characterized first. As shown in Table 1, the distribution of radioactivity among the 5' mononucleotides indicated the adequate degree of purity of all the RNA species, with the exception of RNAs 5 and 7. RNA 5, which was the least abundant RNA species analysed, had slightly more than half of its radioactivity associated with pA, whereas RNA 7 seemed to consist of two components approximately equal in amount, one terminating in pA and the other in pU.

Sequencing of the 5' end proximal segments of the poly(A)-containing RNAs was performed by the methods of Donis-Keller *et al.*¹³ and Simoncsits *et al.*¹⁴, using cleavage of the RNAs by RNases T₁ (G-specific), A (C+U-specific) and U₂ (A-specific), followed by electrophoresis on polyacrylamide/urea gels⁸. The essential features of these RNA sequencing methods, which had been previously reported^{7,8,13,14}, were also observed here (Figs 2, 3, 4). RNases T₁ and U₂ were found to cleave regularly and exclusively after G and A, although with varying efficiency, probably depending on the surrounding sequence. RNase A cleaved exclusively after pyrimidines, but, as described previously, failed to cleave after some, especially when these were part of pyrimidine stretches. Accordingly, the absence of a band in any given position, after digestion with RNase A or U₂ or T₁, was tentatively interpreted as reflecting the presence at that position of a pyrimidine if the data indicated that this putative pyrimidine was part of a stretch of pyrimidines (two or more); in such a situation, the slowest moving among the pyrimidine-terminated oligonucleotides corresponding to the pyrimidine stretch was always clearly visible, with one or more of the other products being sometimes visible as fainter bands.

In most cases, a single band was observed at any given ladder position in the gel track, supporting the idea that a predominant RNA species was being analysed. In some cases, however, a definitely weaker band was observed at the same position as the predominant band, and disregarded as a contaminant (these weaker bands were often also present in the untreated RNA control). In a few cases, no band was seen in any of the lanes at a position corresponding to a ladder step, or two bands of comparable intensity were observed at the same position; in these cases, no conclusive identification was made.

Interpretation of the results was complicated by the presence of extra bands in the gel pattern, especially in the region corresponding to the smaller oligonucleotides. These extra bands were in general less strong than the *bona fide* specific RNase cleavage products. Because of this and because they did not correspond with bands visible in the ladder or corresponded with obviously abnormal steps of the ladder, these extra bands could usually be recognized as spurious. The nature of these extra bands is not clear. They may be due to the presence during the 5'-end-labelling reaction of contaminating RNA species or DNA fragments (deriving from DNA degradation during the DNase treatment step) in the RNA preparation, or they could reflect the formation, during the enzymatic or chemical cleavage of the RNA, of oligonucleotides with different phosphate end-groups (2', 3'-cyclic phosphate, 2' and 3' phosphate). The latter explanation seems the most likely to account for the fact that, in some sequencing analyses (RNAs 5, 7 and 17), both the specific RNase cleavage products and the chemical cleavage products gave doublets instead of single bands.

In each experiment, an untreated RNA sample was run in parallel with the RNase or chemically treated samples. Usually this control ('-Enz') did not show bands or exhibited them only in the region of the larger oligonucleotides.

The interpretation of the sequence of RNA 7 (Fig. 4) offered some difficulties because of the repetition of each band at the next position in the ladder. A similar observation had been made

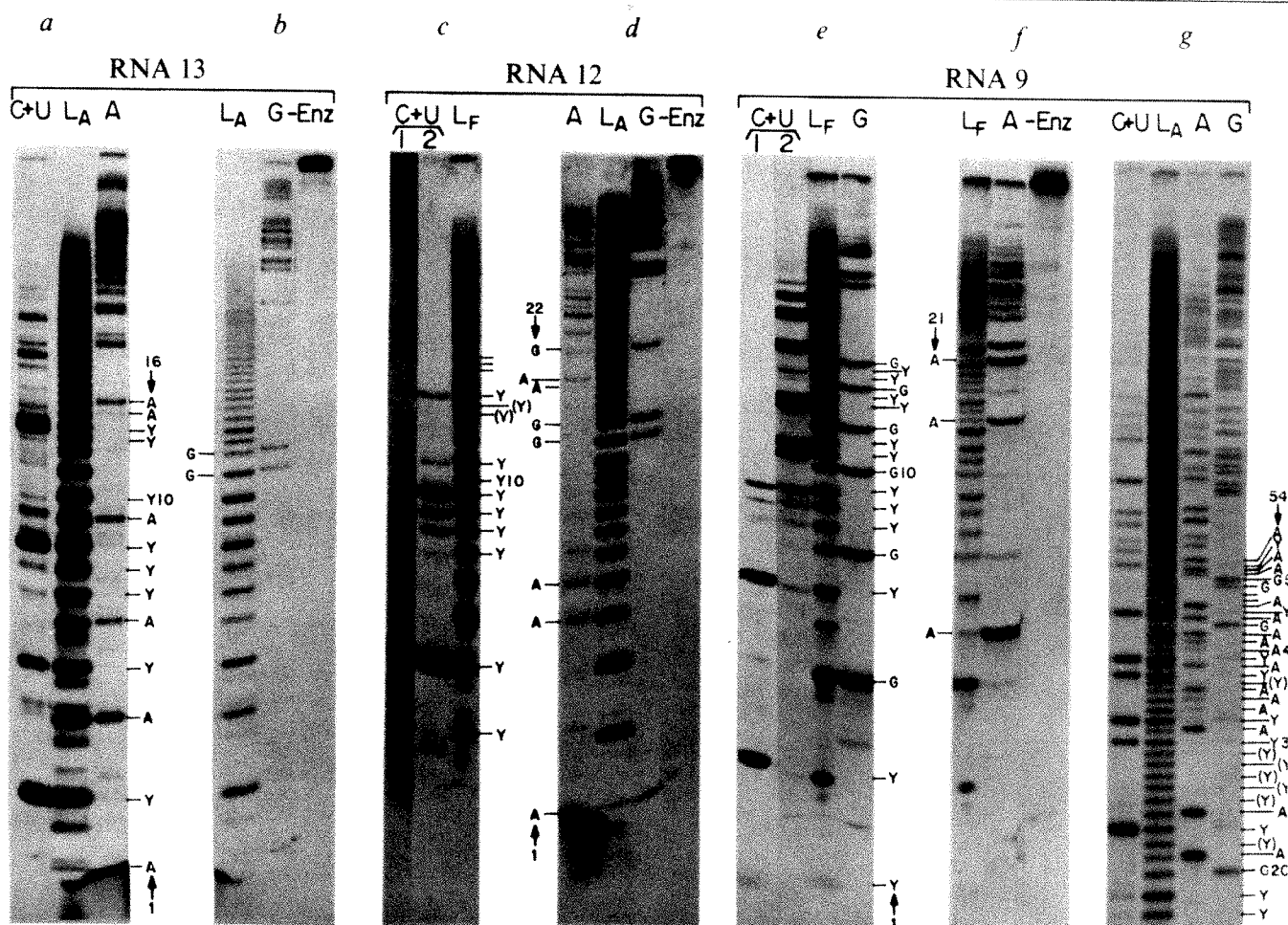


Fig. 2 Sequencing gels of mitochondrial poly(A)-containing RNA species 13, 12 and 9. *a, b*, RNA 13, 25% polyacrylamide/urea gel. Shown at the top of the gel are the cleavage specificities: RNase A (C+U), alkaline ladder (L_A), RNase U₂ (A), RNase T₁ (G) and RNA incubated without enzyme (-Enz). Y indicates a pyrimidine band. *c, d*, RNA 12, digested and analysed as indicated for RNA 13. The two left lanes in *c* show the digests obtained with decreasing amounts of RNase A. L_F indicates a formamide ladder. The horizontal ticks without a letter designation indicate ladder positions for which no nucleotide assignment has been made. *e-g*, RNA 9, digested as indicated for RNA 13 and analysed on a 25% (*e, f*) or, for the sequencing of residues from ~20 to 60, a 10% (*g*) polyacrylamide/urea gel. Details of the base-specific enzymatic cleavage reactions^{7,13,14} and the conditions of electrophoretic analysis^{7,8} are described elsewhere. In some experiments, to visualize a pyrimidine band in the first or second position of the sequence, the amount of RNase A used (per ~30 ng RNA and 3 µg yeast tRNA carrier) was increased to 20 ng.

previously for 12S rRNA⁸ and interpreted to reflect the existence of two species of 12S rRNA differing in molecular size by one nucleotide at the 5' end. With RNA 7, in agreement with the P₁ digestion data (Table 1), analysis of the sequencing gel showed that it could be interpreted as resulting from the electrophoretic separation of the products of RNase digestion of two species, one starting with AUG and the other with UG; comparison with the DNA sequence (see below) confirmed the validity of this interpretation. RNA 5 is present in HeLa cells in minute amounts, and only a partial purification could be obtained even after two electrophoretic runs through CH₃HgOH-agarose slab gels, as revealed by the P₁ digestion data (Table 1). Therefore, only partial sequence data could be obtained for this RNA species. However, we could identify with reasonable confidence the first two nucleotides (AY), the sixth (Y) and the ninth (G) (not shown).

Most mRNAs start at or very near an AUG or AUA triplet

Using the criteria described above, a stretch of 10–54 nucleotides, depending on the RNA species, could be read in the sequence of all the poly(A)-containing RNAs analysed here (except RNA 5). The sequences thus obtained are shown in Fig. 5; the corresponding DNA sequences (refs 4, 15 and 26 (accompanying paper)) are also reported to show their perfect alignment with the RNA sequences. Comparison of the RNA

and DNA sequences clearly indicates that all the AYG and AYA triplets underlined in the RNA sequences are indeed AUG and AUA codons. The partial sequence data for RNA 5 are consistent with the first triplet being an AUA. Thus, the striking results of this analysis are that, with one possible exception (RNA 12), all the putative human mitochondrial mRNAs start directly with an AUG or AUA triplet (which is a methionine codon in human mitochondria¹) or have a few nucleotides (1–3) preceding the first AUG or AUA. Sequencing data of human¹⁶ and bovine cytochrome *c* oxidase subunits I (COI) and II^{1,4} and comparison with the sequences of the yeast cytochrome *b* (A. Tzagoloff, personal communication) and cytochrome *c* oxidase subunit III (COIII)¹⁷ genes have indicated that the first AUGs of RNA 9 (COI mRNA), RNA 16 (COII mRNA), RNA 15 (COIII mRNA) and RNA 11 (cytochrome *b* mRNA) are initiator codons for the corresponding polypeptides. Thus, it seems reasonable to extrapolate from these results and interpret the 5'-proximal AUGs of the other mRNAs likewise as initiator codons. The mtDNA sequencing data suggest that AUA may function as an initiator codon in human mitochondria¹, and the data presented here support this. If the first AUA of RNA 13 is indeed the initiator codon, then, with the exception of RNA 12, all putative mitochondrial mRNAs have the initiator codon within the first six nucleotides. RNA 12 may also follow this rule. In fact, in bovine mtDNA, the first triplet of the reading frame corresponding to RNA 12 is AUA, instead of AUU as in human mtDNA; this observation sugges-

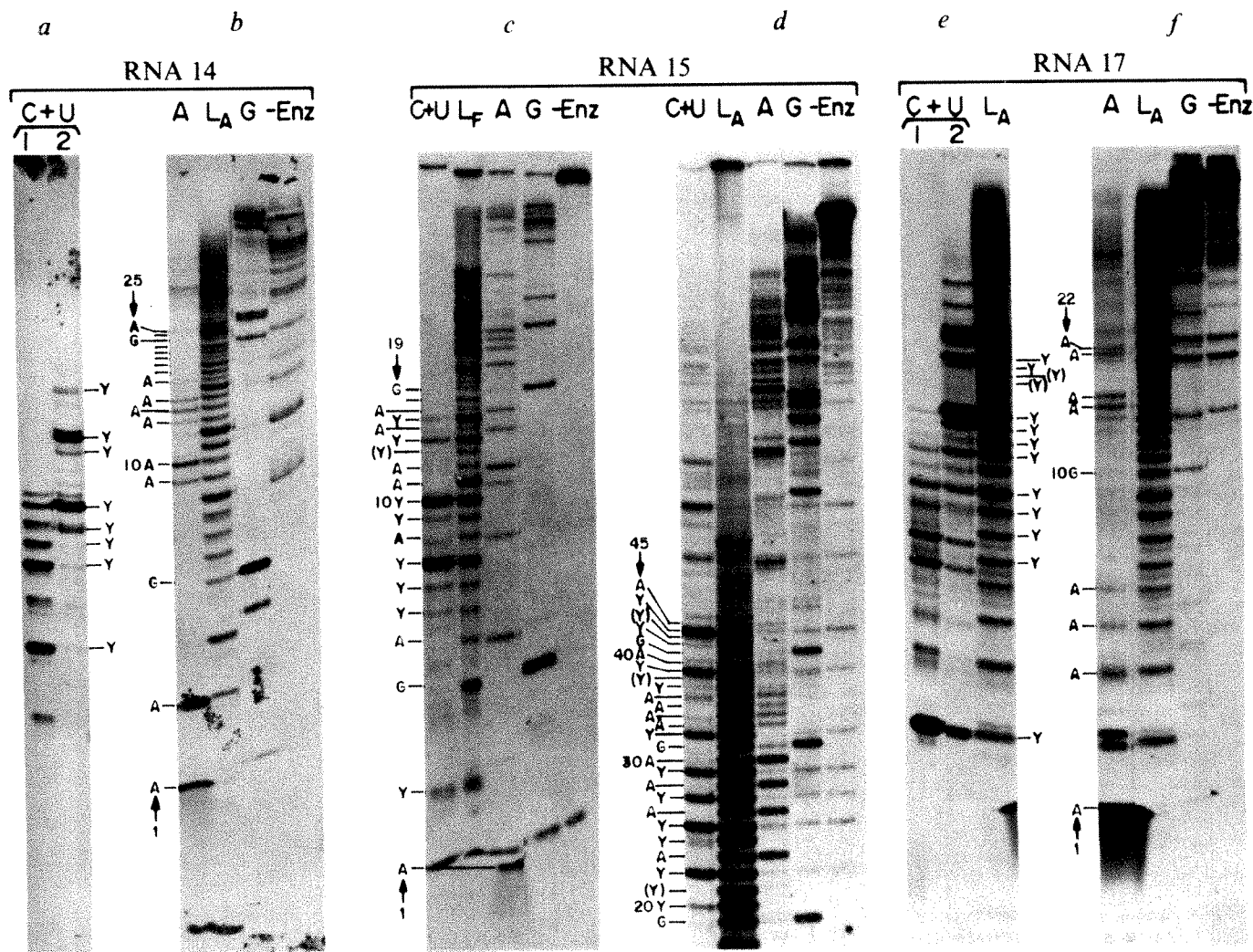


Fig. 3 Sequencing gels of mitochondrial poly(A)-containing RNA species 14, 15 and 17. *a, b*, RNA 14, 25% polyacrylamide/urea gel. *c, d*, RNA 15, digested as shown for RNA 14, and analysed on a 25% (*c*) or a 10% (*d*) polyacrylamide/urea gel. *e, f*, RNA 17, digested and analysed as described for RNA 14.

ted that AUU may function as initiator codon for the polypeptide coded for by RNA 12 (ref. 1). If the 5'-end heterogeneity observed in RNA 7 is an *in vivo* phenomenon, and not a consequence of degradation during the preparation and *in vitro* labelling of the RNA, and if the molecules starting with UG also function as mRNAs, the first initiator codon which could be used in these molecules would be the AUA occurring 14-nucleotides downstream from the 5' end.

How do ribosomes attach to mitochondrial mRNAs?

The finding that most mitochondrial mRNAs either start directly at the initiator codon or have only a few nucleotides (1–8) preceding this codon poses interesting questions concerning the mechanism whereby mitochondrial ribosomes attach themselves to these messengers. Both in eukaryotic and prokaryotic mRNAs, there is a stretch of variable length preceding the initiator codon^{5,6}. In prokaryotic mRNAs there is good evidence that this stretch contains a ribosome-binding site, which includes a properly positioned sequence complementary to the 3' end of 16S rRNA^{18,19}. A single exception to this rule is that of the λ repressor mRNA involved in repressor maintenance, which starts directly at the translation initiator codon²⁰. In this case, the low efficiency of translation of the mRNA was attributed to the lack of a strong ribosome-binding site. The observation that, in the 5' non-coding region of some eukaryotic mRNAs, there is considerable complementarity with a purine-

rich sequence near the 3' end of 18S rRNA⁵ has suggested that specific mRNA-rRNA base pairing may also help the positioning of eukaryotic ribosomes on the mRNAs; however, the variability in the position of the pyrimidine-rich sequence within the mRNA molecules and even the lack of these sequences in some mRNAs questions the generality of this mechanism⁶.

Sequencing analysis of a variety of eukaryotic (cellular and viral) mRNAs has shown that in most, if not all these mRNAs, translation starts at the AUG closest to the 5' end, apparently independently of the length of the 5' non-coding region (varying in the known cases between 11 and over 200 nucleotides) and of its nucleotide sequence^{6,21}. No minimum distance between the initiator codon and 5' end of the mRNA, which would allow proper initiation, has yet been defined for cellular or viral mRNAs; however, in the case of Sindbis virus RNA, there is evidence suggesting that the AUG immediately adjacent to the cap²² is not used for initiation⁶; this would imply that cytoplasmic ribosomes cannot recognize an initiator codon in such a location. To explain why initiation of translation is limited to the most 5'-proximal AUG, a 'scanning' mechanism has been proposed⁶, whereby a 40S ribosomal subunit would attach initially to the 5' end of an RNA chain and subsequently migrate towards the interior of the RNA, stopping when the first AUG is encountered; at this point the 60S subunit would join and translation start. In this model, the cap structure would enhance the binding of the 40S subunits to the 5' end of the mRNA, but would not be required. It is interesting that HeLa cell, and in

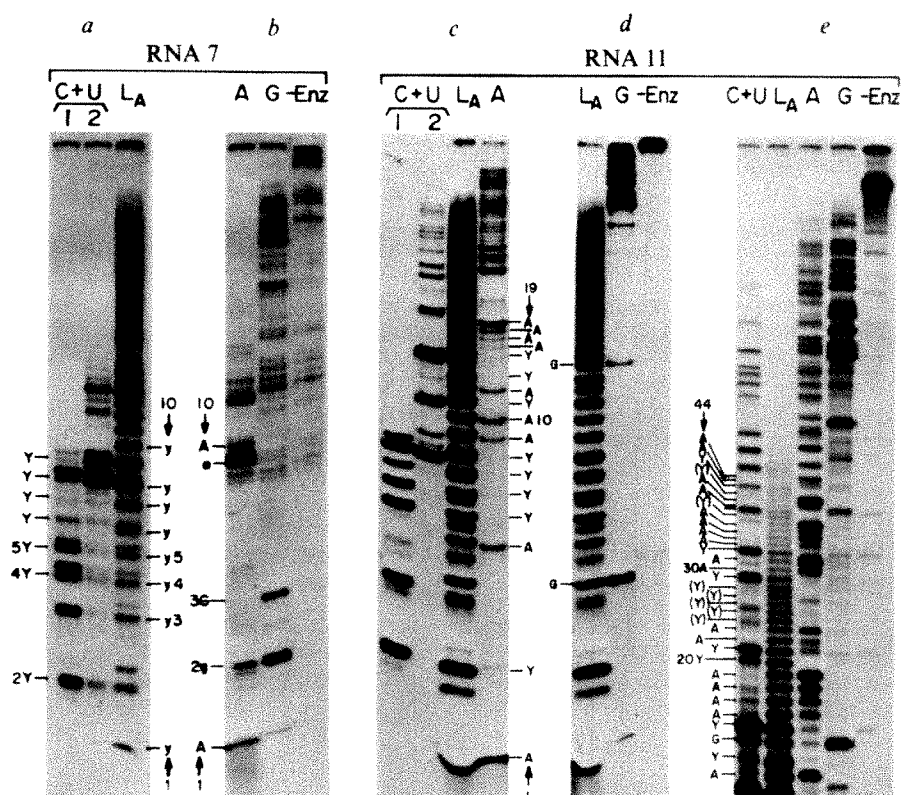


Fig. 4 Sequencing gels of mitochondrial poly(A)-containing RNA species 7 and 11. *a, b*, RNA 7, 25% polyacrylamide/urea gel. The capital and lower case letters indicate, respectively, the nucleotide assignments of the presumptively longer RNA species and of a species differing from this one by the lack of the 5'-terminal nucleotide at the time of labelling. *c-e*, RNA 11, digested as indicated for RNA 7, and analysed on a 25% (*c, d*) or a 10% (*e*) polyacrylamide/urea gel.

RNA 5

5' AY---Y--G.....
1
tRNA^{Leu}
ATAAAAGTAATAACCATG.....
1

RNA 7

5' AYGYYYYYYA.....
1 10
tRNA^{Arg}
ATTTACCAATATGCCCTCA.....
1 10

RNA 9

5' YYGAYGYYYGYGGAYYGYGAYYAYYNYNYAYAAAAAYYAYAAAGAYA--GGAAYA.....
1 10 20 30 40
CTGATGTTTCGCCGACCGTTGACTATTCTCTACAAACCCACAAGACATTGGAACA.....
1 10 20 30 40

RNA 11

5' AYGAYYYYAAYAGYAAAAAYYAAIYNYNYNYAYAAAAAYYAAIYIYAA.....
1 10 20 30 40
ATGACCCCAATACGCAAAATTAACCCCTAATAAAATTAATTAA.....
1 10 20 30 40

RNA 12

5' AYYAAYYYYGGYNYIAA.....
1 10
tRNA^{F-Met}
CCCGTACTAATTAATCCCTGGCCCA.....
1 10

RNA 13

5' AYAYAYYYYAYGGYYAA.....
1 10
tRNA^{Leu}
TTCTTAACAACATACCCATGGCCAA.....
1 10

RNA 14

5' AAYGYYYYAAYYAAAAA.....
1 10
tRNA^{Lys}
TTACAGTGAATATGCCCAACTAAATA.....
1 10

RNA 15

5' AYGAYYYAYYAAIYIYAYA-GYIYIYAYYAYAGYAAAAAYIYAGYIYIYA.....
1 20 30 40
ATGACCCACCAATCACATGCCCTATCATATAGTAAACCCAGCCCA.....
1 10 20 30 40

RNA 16

5' AYGGYAYAYGYAGYGAAGYAGGYYAYA.....
1 10 20
tRNA^{Asp}
TATATCTTAATGGCACATGCAGCGCAAGTAGGTCTACA.....
1 10 20

RNA 17

5' AYAAYYYYGYYYAAIYNYIAA.....
1 10 20
tRNA^{Gly}
AAAAGAGTAATAAACTTCGCCTTAATTTTAA.....
1 10 20

Fig. 5 The sequences of the 5'-end proximal regions of the mitochondrial poly(A)-containing RNAs determined in the present work are shown, aligned with the corresponding mtDNA sequences. A 3'-proximal segment of the sequence of the tRNA genes lying on the 5' side of most of the poly(A)-containing RNA coding stretches is also shown. Y indicates a pyrimidine (the parentheses denote a tentative identification, as explained in the text), a dash (—), an unidentified base; at position 38 in the RNA 11 sequence two possible assignments are indicated.

general, presumably mammalian, mitochondrial mRNAs are not capped²³.

The terminal or subterminal position of the initiator codon in human mitochondrial mRNAs may eliminate the need for a scanning process, at least an extensive one. A plausible mechanism would involve attachment of the ribosome at or near the 5' end of the mRNA with recognition of the initiator codon either directly or after a fine adjustment. The secondary structure of mitochondrial mRNAs may be such that it would exclude all internal sites containing AUG or AUA codons, exposing only the terminal or subterminal one. It is also conceivable that the special features of mammalian mitochondrial ribosomes or some initiation factor would make the ribosomes suitable for recognizing the terminal or subterminal initiator codons. Analysis of the potential secondary structures of human mitochondrial mRNAs and binding studies with mitochondrial or other ribosomes should elucidate the mechanisms operating in the initiation of translation in human mitochondria.

The 5' ends of most mRNA coding sequences are immediately adjacent to tRNA genes

The complete absence or minimal length of the 5' non-coding stretch in human mitochondrial mRNAs may represent the vestige of a primitive organization, or may reflect, with other traits (for example, the relatively small size of the rRNA and tRNA species), an evolutionary trend to simplicity and economy

of the mitochondrial genome²⁴. In any case, the degree of proximity of the initiator codon to the 5' end of the mitochondrial mRNAs seems to be dictated primarily by the features of gene organization of this genome, in particular, by the positions of the tRNA genes. In fact, whatever the location of the initiator codon relative to the 5' end of the mitochondrial mRNAs, the data presented here (Fig. 5) suggest that whenever a tRNA gene exists on the 5' side of the mRNA coding sequence, the latter starts immediately after the tRNA gene. This observation confirms and refines the results of a recent detailed transcription analysis of HeLa cell mtDNA³. These mapping and sequencing data strongly support the idea that the tRNA sequences represent the recognition signals for a putative processing enzyme which, by precise endonucleolytic cleavages, would release the 5' ends of the mature mRNAs^{2,3}. As will be discussed elsewhere²⁵, recent experimental evidence indicates that the tRNA sequences may perform a similar function in the processing steps which release the 3' ends of the human mitochondrial mRNAs. These striking features of the organization of the human mitochondrial genome underlie the crucial part that the tRNA sequences probably play in mitochondrial RNA processing in human cells^{2,3}.

This work was supported by NIH grants GM-11726 and GM-07616. We thank Drs F. Sanger and B. Barrell for communicating the human mtDNA sequence before publication. The technical assistance of Ms A. Drew is gratefully acknowledged.

Received 1 September; accepted 27 November 1980.

1. Barrell, B. G. *et al.* in *31st Moshbach Colloquium on Biological Chemistry of Organelle Formation* (eds Weiss, H. & Sebald, W.) (Springer, Berlin, in the press).
2. Attardi, G. *et al.* in *The Organization and Expression of the Mitochondrial Genome* (eds Kroon, A. M. & Saccone, C.) 103–119 (North-Holland, Amsterdam, 1980).
3. Ojala, D., Merkel, C., Gelfand, R. & Attardi, G. *Cell* **22**, 393–403 (1980).
4. Barrell, B. G., Bankier, A. T. & Drouin, J. *Nature* **282**, 189–194 (1979).
5. Hagenbüchle, O., Santer, M., Steitz, J. A. & Mans, R. J. *Cell* **13**, 551–563 (1978).
6. Kozak, M. *Cell* **15**, 1109–1123 (1978).
7. Ojala, D., Montoya, J. & Attardi, G. *Nature* **287**, 79–82 (1980).
8. Crews, S. & Attardi, G. *Cell* **19**, 775–784 (1980).
9. Amalric, F., Merkel, C., Gelfand, R. & Attardi, G. *J. molec. Biol.* **118**, 1–25 (1978).
10. Attardi, G. *et al.* in *Proc. ICN-UCLA Symp. molec. cell. Biol.* **15**, 443–469 (1979).
11. Gelfand, R. thesis, Calif. Inst. Technol. (1980).
12. Fujimoto, M., Kuninaka, A. & Yoshino, H. *Agric. biol. Chem.* **38**, 1555–1561 (1974).

13. Donis-Keller, H., Maxam, A. M. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527–2538 (1977).
14. Simoncits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. & Guillely, H. *Nature* **269**, 833–836 (1977).
15. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).
16. Chomyn, A., Hunkapiller, M. & Attardi, G. *Nucleic Acids Res.* (in the press).
17. Thalenfeld, B. & Tzagoloff, A. *J. biol. Chem.* **255**, 6173–6180 (1980).
18. Shine, J. & Dalgarno, L. *Biochem. J.* **141**, 609–615 (1974).
19. Steitz, J. A. & Jakes, K. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4734–4738 (1975).
20. Ptashne, M. *et al. Science* **194**, 156–161 (1976).
21. Sherman, F., Stewart, J. W. & Schweingruber, A. M. *Cell* **20**, 215–222 (1980).
22. Hefti, E., Bishop, D. H. L., Dubin, D. T. & Stollar, V. J. *Virology* **17**, 149–159 (1976).
23. Grohmann, K., Amalric, F., Crews, S. & Attardi, G. *Nucleic Acids Res.* **5**, 637–651 (1978).
24. Attardi, G. *Trends biochem. Sci.* (in the press).
25. Ojala, D., Montoya, J. & Attardi, G. *Nature* **290**, 470–474 (1981).
26. Anderson, S. *et al. Nature* **290**, 457–465 (1981).

tRNA punctuation model of RNA processing in human mitochondria

Deanna Ojala, Julio Montoya & Giuseppe Attardi

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

A 3'-end proximal segment of most of the putative mRNAs encoded in the heavy strand of HeLa cell mtDNA has been partially sequenced and aligned with the DNA sequence. In all cases, the 3'-end nucleotide of the individual mRNA coding sequences has been found to be immediately contiguous to a tRNA gene or another mRNA coding sequence. These and previous results indicate that the heavy (H) strand sequences coding for the rRNA, poly(A)-containing RNA and tRNA species form a continuum extending over almost the entire length of this strand. We propose that the H strand is transcribed into a single polycistronic RNA molecule, which is processed later into mature species by precise endonucleolytic cleavages which occur, in most cases, immediately before and after a tRNA sequence.

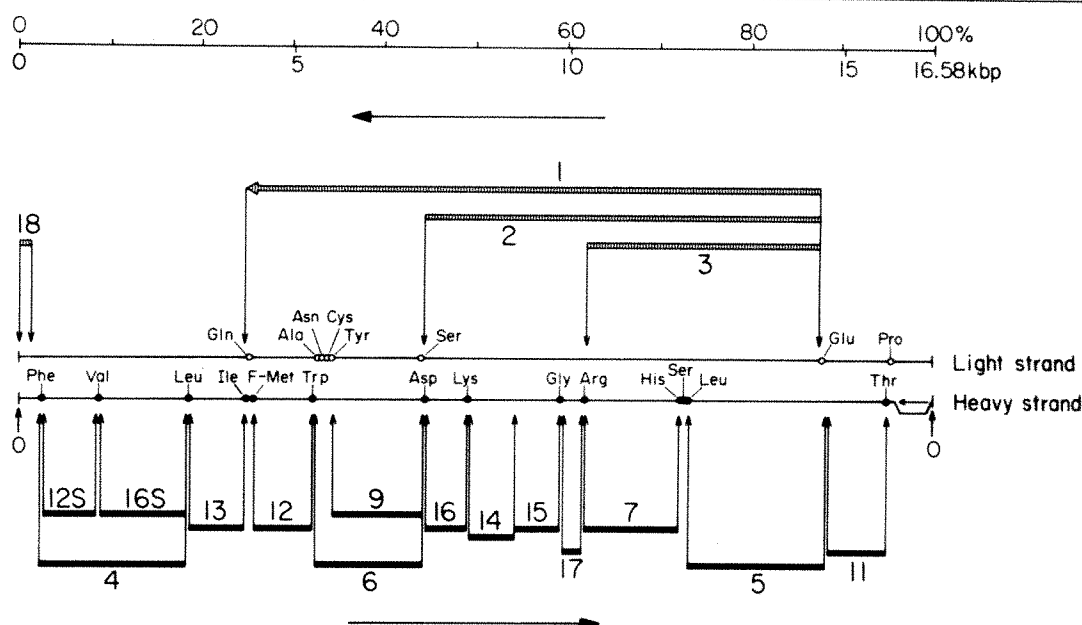
RECENT transcription mapping analysis of HeLa cell mitochondrial DNA (mtDNA) has revealed that the H-strand sequences specifying the rRNA and poly(A)-containing RNA species (which are probably specific mRNAs or their precursors) are flanked on one side, and often on both sides, by a tRNA gene^{1,2} (Fig. 1). In agreement with these observations, the alignment with the mtDNA sequence (refs 3–5 and 37) of a 5'-end proximal sequence of all H strand-coded putative mRNAs has shown that the sequences complementary to these mRNAs are almost always immediately adjacent, without any intervening nucleotide, to a tRNA gene or an mRNA coding sequence^{6,7}.

Here, we report that each of the eight mRNAs analysed is immediately contiguous, on its 3' side, to a tRNA gene or an mRNA coding sequence. On the basis of these and previous results^{2,4,6–8}, we propose a model of H-strand gene expression which involves transcription of almost the entire length of this strand in the form of a single polycistronic RNA molecule and its subsequent processing directed by the tRNA sequences.

Isolation and 3'-end sequencing analysis of mitochondrial mRNAs

The putative mRNAs were isolated from the poly(A)-containing RNA fraction of micrococcal nuclease-treated mitochondria

Fig. 1 Transcription map of HeLa cell mtDNA. The two mtDNA strands have been linearized at the origin of replication (0). The position and identities of the tRNA genes on the H strand (●) and on the L strand (○) were sequence (refs 16 and 37). The solid and hatched bars indicate the H-strand and L-strand transcripts, respectively; numbers refer to the various poly(A)-containing RNA species according to the classification of Amalric *et al.*¹⁹. The upper and lower arrows indicate the direction of L- and H-strand transcription, respectively. kbp, Kilobase pairs.



by electrophoresis through an agarose-CH₃HgOH slab gel, elution of the individual species from the gel, treatment with RNase-free DNase and re-run of each species on a separate gel of the same type, as described elsewhere⁶.

Partial sequence analysis of the mRNAs was carried out by an adaptation of the minus sequencing method using 'phased' oligo (dT) primers and reverse transcriptase^{9,10}, as described in Fig. 2 legend. In the transcriptase reaction, labelled primers and unlabelled dNTPs¹¹ were used rather than unlabelled primers and [α -³²P]dNTPs, as in the usual procedures; in the latter approach, the probable presence of contaminating oligonucleotides resulting from the DNase treatment (which could function as primers) was expected to complicate considerably the interpretation of the results.

To determine the correct primer to be used with the RNA being investigated (and thus the identity of the nucleotide adjacent to the poly(A) tail) each of three 5'-end-labelled primers [p(dT)₈-dA, p(dT)₈-dC and p(dT)₈-dG] was tested for its ability to prime cDNA synthesis using the RNA as a template. As shown previously^{9,10}, the presence of a specific nucleotide at the 3' end of the oligo(dT) stretch ensured that only the primer molecules hybridized with the RNA in phase (that is, those base-paired with the eight most 5'-proximal residues of the poly(A) tail and the first nucleotide on the 5' side of this tail) would promote extension synthesis by the reverse transcriptase. The band pattern formed by the products of this synthesis showed a highly variable intensity of the individual bands, and was very typical and highly reproducible for each RNA, thus providing a diagnostic test of a positive result; only limited extension synthesis was observed when the two other, non-specific primers were tested with the same RNA.

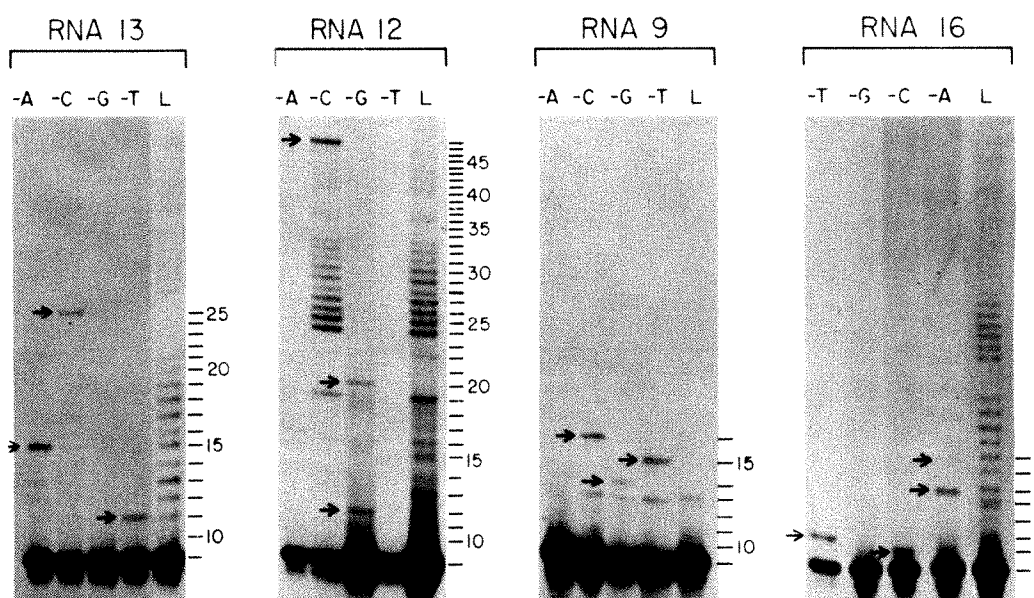
The correct oligo(dT)-dN primer identified for each RNA was incubated with the corresponding RNA, dNTPs and reverse transcriptase in four incubation mixtures, each lacking one of the four dNTPs. Figure 2 shows representative results obtained with some of the various putative mRNAs analysed. In general, in each 'minus' reaction mixture, a prominent band can be observed which corresponds to a specific ladder position. This band represents the oligo(dT) primer extended by the reverse transcriptase up to the position preceding, in the cDNA sequence, the nucleotide missing from the reaction mixture. For example, in the experiment using RNA 13, the '-A' lane shows a band at ladder position 15, the '-C' lane a band at position 25, and the '-T' lane a band at position 11. Therefore, the identities of the nucleotides at positions 16, 26 and 12 in the cDNA can be interpreted as A, C and T, respectively, and those of the

corresponding nucleotides in the RNA template U, G and A. The '-G' lane shows no band above the primer, which indicates the presence of a C residue at the position in the RNA sequence corresponding to step 10 of the ladder. The other RNA species gave results similar to those obtained for RNA 13. However, in some cases, an unambiguous identification of the stop positions corresponding to all four nucleotides could not be made. In particular, RNA 12 did not produce the expected band at position 10. Recent experiments suggest that this is probably due to the melting of the short extended primer (10 nucleotides long, comprising Ts and As) from the RNA template during the chromatography step. In the case of RNA 7, the strong signal from the band of the primer (position 9) prevented the recognition of the band at position 10. In these cases, a positive identification could not be made for position 10 or 11 in the cDNA. A faint second band was observed in some cases at a position higher up than that of the first band (see the '-G' lane for RNA 12 and the '-A' lane for RNA 16). This second band was tentatively interpreted to be due to partial readthrough by the transcriptase past the first stop position (presumably because of the presence of contaminating nucleotides in the reaction mixture) and arrest at the next stop position dictated by the sequence. In some experiments, multiple prominent bands were observed in an individual lane at positions below the stop band; these generally correlated with prominent ladder bands (for example, the '-C' lane for RNA 12), and were interpreted as reflecting pauses in the transcriptase progression, as previously observed¹², and were thus disregarded.

mRNA coding sequences are immediately adjacent at their 3' end to tRNA genes or other mRNA coding sequences

The partial sequencing data determined for the 3'-end proximal segments of the putative mRNAs are shown in Fig. 3. The corresponding DNA sequences (refs 3-5 and 37) are also reported to show their alignment with the nucleotides identified in the RNAs. Our experimental data allowed unambiguous matching of the RNA and DNA sequences. The only discrepancy observed is at a residue in RNA 11 corresponding to ladder position 13, where A was found instead of G. This may reflect a sequence difference between human placenta mtDNA (used for DNA sequencing) and HeLa mtDNA; such sequence differences have been previously observed at other positions⁴. It is interesting that the base substitution mentioned above does not change the codon assignment at that position (tryptophan). It is clear from

Fig. 2 Reverse transcriptase products in reactions using 'phased' oligo(dT) primers and poly(A)-containing RNAs 13, 12, 9 and 16. Partial sequence analysis of the 3'-end proximal segment of each RNA species was done by an adaptation of the 'minus' sequencing method using 'phased' oligo(dT) primers and reverse transcriptase^{9,10}, which will be described in detail elsewhere. Briefly, the p(dT)₈-dA, p(dT)₈-dC and p(dT)₈-dG primers were 5'-end labelled with [γ -³²P]ATP (8,000 Ci mmol⁻¹) and T4 polynucleotide kinase, after dephosphorylation with bacterial alkaline phosphatase⁸, then run on a



15% polyacrylamide slab gel in Tris-borate-EDTA. After autoradiography, the band corresponding to nonanucleotides in each lane, which contained most of the radioactivity, was eluted; the final specific activity of the labelled primers was $5-10 \times 10^7$ c.p.m. μg^{-1} . To determine the nucleotide adjacent to the poly(A) tail in the RNAs to be analysed, each RNA species (0.05–0.10 pmol) was mixed at 4 °C with a 100-fold molar excess of each of the 5'-end-labelled primers and 50 mM of each of the four dNTPs in 20 μl of 0.05 M Tris, pH 8.0 (25 °C), 0.05 M KCl, 0.005 M MgCl₂ and 0.01 M dithiothreitol. After addition of AMV reverse transcriptase (4 U), the reaction mixtures were incubated for 3 min at 39 °C. In some experiments, to facilitate the hybridization of the primer with the RNA, the mixtures were incubated for 5 min at 75 °C and then cooled for 15 min at 4 °C, before addition of the reverse transcriptase; however, no difference in results was observed. After stopping the reaction by cooling and addition of 0.01 M EDTA, the reaction products were precipitated with ethanol in the presence of 0.3 M NaCl and carrier yeast tRNA, the precipitates washed with ethanol, dissolved in 0.001 M Tris-EDTA, mixed with an equal volume of 10 M urea, heated for 2 min at 90 °C, then run on a thin (0.5 mm) 10% polyacrylamide/7 M urea sequencing gel for 3.5 h at 1,000 V. The correct oligo(dT) primer for each RNA, identified from the electrophoretic pattern of the products (see text: p(dT)₈-dA for RNAs 13 and 12, p(dT)₈-dG for RNAs 9 and 16), was incubated as described above with the corresponding RNA, dNTPs and reverse transcriptase in four reaction mixtures, each lacking one of the four dNTPs. In addition, a 'ladder' (L) reaction was performed in which all the four dNTPs were present. After stopping the reaction by quick cooling and addition of EDTA, excess primer was separated from the RNA-DNA hybrid by passage through a Sephadex G50 column equilibrated with 0.3 M NaCl, 0.01 M Tris, pH 7.4, in the presence of yeast tRNA carrier. The material in the void volume was ethanol-precipitated and then run on a 10% polyacrylamide/7 M urea sequencing gel. Arrows indicate the bands interpreted as corresponding to stop positions in the reverse transcriptase reaction. The numbering of the ladder steps indicates the molecular length (number of nucleotides) of the extended primers, starting from the 5'-end nucleotide.

Fig. 3 that the most 3'-proximal nucleotide identified directly in the RNA sequence either corresponds in the DNA to a residue immediately adjacent to the 5' end of a tRNA gene (RNAs 12, 16, 15, 7 and 11), or is separated from it by one to three As (RNAs 13 and 9). The most 3'-end proximal nucleotide identified in RNA 14 corresponds in the DNA to a residue separated by one A from the initiator codon (AUG) of the flanking RNA 15 coding sequence. Our experimental approach to the 3'-end sequence analysis did not allow discrimination between any 3'-terminal A encoded in the DNA and the As of the poly(A) tail. However, as in most of the mRNAs analysed, the most 3'-proximal nucleotide different from A corresponds to a nucleotide in the DNA immediately contiguous to a tRNA gene, it seems reasonable to extrapolate from these results and to assume that, in the other mRNA also, where one to three As separate this nucleotide from a tRNA gene or an mRNA coding sequence, these As are similarly transcribed from the DNA.

Comparison of the 3'-end proximal sequence of the small mitochondrial rRNA from hamster cells¹³ with the sequence of the corresponding segment of the human 12S rRNA gene⁴ shows a high degree of homology, with the most 3'-proximal nucleotide of the hamster RNA missing in the human mtDNA sequence which just precedes the tRNA^{Val} gene (Fig. 3). Therefore, it seems very likely that, as in mouse mtDNA¹⁴, the human 12S rRNA gene extends to the tRNA^{Val} gene, thus following the same rule described for the mRNA coding sequences. No direct information about the 3'-end proximal sequence of human 16S rRNA is available; however, comparison of the length of this RNA determined from S₁ protection data (~1,600 nucleotide)¹⁵ with the length of the interval between the tRNA^{Val} and the tRNA^{Leu} genes (1,559 nucleotides)⁴ strongly suggests that the 3'

end of the 16S rRNA corresponds to a residue in DNA immediately adjacent, or very close, to the tRNA^{Leu} gene.

In view of the exceedingly small amount of material available, no 3'-end sequence information was obtained for either RNA 5 or 17. However, the almost perfect correspondence between the molecular size of RNA 17, as estimated from S₁ protection data (340 nucleotides), and the length of the DNA sequence between tRNA^{Gly} and tRNA^{Arg} (346 nucleotides)², and the previously demonstrated juxtaposition of the tRNA^{Gly} gene with the 5' end of the RNA 17 coding sequence⁷ strongly suggest that the 3' end of the sequence specifying RNA 17 is immediately adjacent to the tRNA^{Arg} gene. Similarly, the mapping contiguity of the sequences coding for RNAs 5 and 11 (ref. 2) probably reflects their immediate juxtaposition at the nucleotide level.

Most mitochondrial mRNAs terminate with U or UA

Most of the 3'-end proximal sequences of the mitochondrial mRNAs terminate with U or UA (Fig. 3). The significance of this observation has been clarified by the analysis of the human mtDNA sequence (refs 16 and 37). It has been found that many reading frames, including those corresponding to the mRNAs mentioned above, lack a stop codon, and that, in these cases, a T or TA follows the last sense codon and immediately precedes a tRNA gene or another reading frame. This observation has led to the suggestion that poly(A) addition to the 3'-terminal U or UA of the transcripts of these reading frames may create the missing stop codon. The 3'-end proximal sequences determined here for the mitochondrial mRNAs strongly support this model for polypeptide chain termination in human mitochondria. The

reading frames corresponding to RNAs 9 and 16 have stop codons AGA and UAG, respectively^{3,5} (AGA is probably a stop codon in human mitochondria¹⁷). The observation that in these mRNAs there is a 3' noncoding stretch extending to the 5' end of the flanking tRNA gene indicates that, as in the case of the 5' terminus of the human mitochondrial mRNAs⁷, the rule dictating the possible presence and length of a 3'-end segment flanking the polypeptide coding sequence is the position of the tRNA genes in the DNA sequence.

The tRNA punctuation model of H-strand gene expression

The sequence data discussed above and in the accompanying paper⁷ and the results of previous work^{4,8} indicate that the H-strand sequences coding for the rRNAs, poly(A)-containing RNAs and tRNAs are immediately contiguous to each other, extending continuously from coordinate 2/100 to coordinate 95/100 (relative to the origin 0/100). This arrangement is consistent with a model of transcription of the H strand in the form of a single molecule which is processed by precise endonucleolytic cleavages before and after each tRNA

sequence to yield the mature products or, in some cases, processing intermediates, like the putative precursor of the rRNAs (RNA 4) and the precursor of RNA 9 (ref. 2).

Recent work has indicated a concentration of H-strand nascent transcripts in the quadrant of the mtDNA *Hpa*II map which is adjacent to the origin of replication in the direction of H-strand transcription¹⁸ (Fig. 4a). These results strongly suggest that the region of mtDNA around the origin of replication contains an initiation site for H-strand transcription—such a site may thus represent the promoter of the single transcript postulated here. Furthermore, the failure to detect any giant-size H-strand transcripts either associated with transcription complexes¹⁸ or as discrete species^{2,19}, in contrast to the occurrence of discrete giant-size L-strand transcripts, points to the possibility that the processing of the H-strand transcripts proceeds while they still reside on the mtDNA transcription complexes. These results taken together support a model in which transcription of the H strand starts near the origin of replication and proceeds uninterruptedly at least up to the distal end of the D-loop, while the growing chains are processed to yield the mature rRNAs, poly(A)-containing RNAs and tRNAs (Fig. 4b).

Hamster
small rRNA



Fig. 3 Partial sequencing data determined for the 3'-end regions of the mitochondrial poly(A)-containing RNAs are shown aligned with the corresponding DNA sequences. A 5'-proximal segment of the sequence of the adjacent tRNA genes and of the RNA 15 coding stretch is also shown. The figure also includes the 3'-end proximal sequence of the hamster small mitochondrial rRNA¹³, together with the corresponding human mtDNA sequence and a portion of the contiguous tRNA^{Val} coding sequence. Asterisks indicate the non-homologous nucleotides.

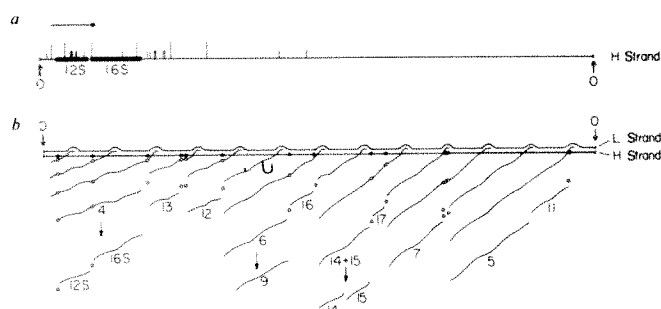


Fig. 4 Proposed model for mtDNA H-strand transcription. *a*, Distribution in the mtDNA H strand of the hybrids formed with nascent RNA chains isolated from transcription complexes (modified from ref. 18). Each tick marks the position in the *Hpa*II map (not shown) of one end of each hybrid, while the other end is arbitrarily localized at the closest, origin-proximal *Hpa*II site; long and short ticks refer to abundant and rare hybrids, respectively. *b*, Diagrammatic representation of the processing of nascent mitochondrial RNA chains in the transcription complexes: ●, tRNA gene; ○, mature tRNA.

In the processing of the primary transcripts, the secondary structure of the tRNA sequences may represent the main recognition signal, providing the punctuation in the reading of mtDNA information¹. In fact, these sequences may acquire a clover-leaf configuration while they are still a part of the nascent transcripts, and the processing enzyme(s) may recognize this structure or a portion of it. It is interesting that the *Escherichia coli* RNase P²⁰⁻²³ and a similar *in vitro* enzymatic activity from *Bombyx mori*²⁴ catalyse, in the tRNA precursors, precise endonucleolytic cleavages at the 5' terminus of the mature tRNAs, while the production of the 3' terminus of mature tRNAs in *E. coli*^{22,25} and *B. mori*²⁴ involves a 3' → 5' exonuclease activity, following an endonucleolytic event at some distance from the 3' end. There is good evidence to indicate that the RNase P recognizes some aspect of the structural conformation of the tRNA rather than the nucleotide sequence at its cleavage sites in the tRNA precursors^{26,27}. It is possible that two enzymatic activities analogous to RNase P, but with different specificities for the 5' and 3' terminus of the tRNAs, are involved in the processing of the H-strand primary transcripts in human mitochondria.

There are a few processing sites in the H-strand transcripts where no tRNA sequences have been found (see Fig. 1)—at such positions, it is conceivable that the processing enzyme(s) recognizes, on one or the other side of the cleavage point, a stem-and-loop structure resembling a portion of a tRNA configuration.

Polyadenylation occurs at endonucleolytic cleavage sites of primary transcripts

One interesting observation is that all the RNA species other than tRNAs which derive from the primary processing of the nascent H-strand transcripts are polyadenylated. This suggests that polyadenylation may be linked in some way to the process-

ing step that releases the 3' end of these RNA products from the flanking tRNA sequences, independently of the functional role of the products. This interpretation may account for the polyadenylation of the presumptive rRNA precursor^{2,28}. The small fraction of 16S rRNA which binds to oligo(dT)-cellulose^{2,19} probably represents incompletely processed 16S rRNA molecules with residual poly(A) tails. It is interesting to note that, in nuclear RNA (both cellular and viral), addition of poly(A) to the 3' ends is a rapid event which seems to precede splicing²⁹⁻³¹; furthermore, at least in most of the transcripts in adenovirus 2 and SV40 infection, the acceptor sites for polyadenylation are generated by endonucleolytic cleavages of larger primary transcripts^{29,32}.

Implications for mitochondrial RNA metabolism

A recent estimation of the steady-state amounts of the mitochondrial mature RNA species has indicated that the molar amount of 12S rRNA is about 60 times higher than that of the most abundant mRNA, RNA 16, with moderate variations in the amounts of the different H strand-coded mRNAs^{33,34}. Because the half life of the 12S rRNA is only 2-5 times longer than that of the mRNAs^{33,34} a difference in the rate of synthesis must be the main factor determining the large difference in amount between the rRNA and mRNA species. In fact, a 50-100-fold higher rate of synthesis has been estimated for 12S rRNA as compared with the mRNAs^{33,34}. Such a difference in rate of synthesis between rRNA and mRNA species could suggest the existence of more than one promoter for H-strand transcription. However, in view of the regular butt-jointing of the sequences coding for all the H-strand transcripts, a more plausible interpretation is that the H strand is in fact transcribed in the form of a single molecule, but that not all transcripts are completed before being released from the template. Premature termination of the transcripts beyond the rRNA cistrons would in fact give a larger molar yield of the rRNA species relative to that of the mRNAs or most of the tRNAs¹. Late in adenovirus 2 infection, a premature termination of transcription, which produces a four- to six-fold greater molar amount of RNA synthesized from the first 2,000 nucleotides of the transcriptional unit than from any other region, has been previously well documented³⁵, and recent evidence has strongly suggested premature termination in heterogeneous nuclear RNA synthesis³⁶. It is attractive to think that the location of the mitochondrial rRNA cistrons close to the initiation point of H-strand transcription in animal cell mtDNA, as has emerged in the course of evolution, may have been advantageous to the cell as one which gave the rRNA cistrons a transcription advantage.

This work was supported by NIH grants GM-11726 and T32 GM-07616. We thank Drs F. Sanger and B. Barrell for communicating the human mtDNA sequence before publication, and Drs D. Dubin and D. A. Clayton for sending us manuscripts in the press. The technical assistance of Ms A. Drew is acknowledged.

Received 28 October 1980; accepted 6 January 1981.

- Attardi, G. *et al.* in *The Organization and Expression of the Mitochondrial Genome* (eds Kroon, A. M. & Saccone, C.) 109-119 (North-Holland, Amsterdam, 1980).
- Cjala, D., Merkel, C., Gelfand, R. & Attardi, G. *Cell* **22**, 393-403 (1980).
- Barrell, B. G., Bankier, A. T. & Drouin, J. *Nature* **282**, 189-194 (1979).
- Eperon, I. C., Anderson, S. & Nierlich, D. P. *Nature* **286**, 450-467 (1980).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
- Cjala, D., Montoya, J. & Attardi, G. *Nature* **287**, 79-82 (1980).
- Montoya, J., Cjala, D. & Attardi, G. *Nature* **290**, 465-470 (1981).
- Crews, S. & Attardi, G. *Cell* **19**, 775-784 (1980).
- Cheng, C. D. *et al.* *J. molec. Biol.* **107**, 527-547 (1976).
- Brownlee, G. G. & Cartwright, E. M. *J. molec. Biol.* **114**, 93-117 (1977).
- Symons, H. R. *Nucleic Acids Res.* **7**, 825-837 (1979).
- Mills, D. R., Dobkin, C. & Kramer, F. R. *Cell* **15**, 541-550 (1978).
- Dubin, D. T. & Baer, R. J. in *The Organization and Expression of the Mitochondrial Genome* (eds Kroon, A. M. & Saccone, D.) (North-Holland, Amsterdam, in the press).
- Van Etten, R. A., Walberg, M. W. & Clayton, D. A. *Cell* **22**, 157-170 (1980).
- Cjala, D. & Attardi, G. *J. molec. Biol.* **138**, 411-420 (1980).
- Barrell, B. G. *et al.* in *31st Moshach Colloquium biol. Chem. Organelle Formation* (eds Weiss, H. & Sebald, W.) (Springer, Berlin, in the press).
- Barrell, B. G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 3164-3166 (1980).
- Cantatore, P. & Attardi, G. *Nucleic Acids Res.* **8**, 2605-2624 (1980).

- Amalric, F., Merkel, C., Gelfand, R. & Attardi, G. *J. molec. Biol.* **118**, 1-25 (1978).
- Fukada, K. & Abelson, J. *J. molec. Biol.* **139**, 377-391 (1980).
- Robertson, H., Altman, S. & Smith, J. J. *biol. Chem.* **247**, 5243-5251 (1952).
- Schedl, P., Primakoff, P. & Roberts, J. *Brookhaven Symp. Biol.* **26**, 53-76 (1974).
- Stark, B. C., Kole, R., Bowman, E. J. & Altman, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3717-3721 (1978).
- Garber, R. L. & Altman, S. *Cell* **17**, 389-397 (1979).
- Sakano, H., Yamada, S., Ikemura, T., Shimura, Y. & Ozeki, H. *Nucleic Acids Res.* **1**, 355-371 (1979).
- Bothwell, A. L. M., Stark, B. C. & Altman, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1912-1916 (1976).
- Schmidt, F. J., Seidman, J. G. & Bock, R. M. *J. biol. Chem.* **251**, 2440-2445 (1976).
- Cleaves, G. R., Jones, T. & Dubin, D. T. *Archs Biochem. Biophys.* **175**, 303-311 (1976).
- Nevins, J. R. & Darnell, J. E. Jr *Cell* **15**, 1477-1493 (1978).
- Schibler, U., Marcu, K. B. & Perry, R. P. *Cell* **15**, 1495-1509 (1978).
- Lai, C. J., Dhar, R. & Khoury, G. *Cell* **14**, 971-982 (1978).
- Fraser, N. W., Nevins, J. R., Ziff, E. & Darnell, J. E. Jr *J. molec. Biol.* **129**, 643-656 (1979).
- Gelfand, R. A. & Attardi, G. (submitted).
- Attardi, G. *et al.* *2nd Int. Cong. Cell Biol.*, Berlin (Springer, Berlin, in the press).
- Fraser, N. W., Sehgal, P. B. & Darnell, J. E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2571-2575 (1979).
- Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S. & Darnell, J. E. Jr *Cell* **19**, 69-78 (1980).
- Anderson, S. *et al.* *Nature* **290**, 457-464 (1981).

Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis

William S. Hayward*, Benjamin G. Neel* & Susan M. Astrin†

* The Rockefeller University, 1230 York Avenue, New York 10021, USA

† The Institute for Cancer Research, Philadelphia, Pennsylvania 19111, USA

Analyses of DNA and RNA from avian leukosis virus (ALV)-induced lymphomas have provided strong evidence that, in most tumours, ALV induces neoplastic disease by activating the c-myc gene, the cellular counterpart of the transforming gene of MC29 virus. The data indicate that, as a rare event, the ALV provirus integrates adjacent to the c-myc gene and that transcription, initiating from a viral promoter, causes enhanced expression of c-myc, leading to neoplastic transformation.

RNA TUMOUR VIRUSES include acute transforming viruses (sarcoma and acute leukaemia viruses) and slow transforming viruses (for example, ALVs and murine leukaemia viruses). The mechanism by which the slow transforming viruses induce neoplastic disease has, until recently, remained obscure. These viruses differ from the acute transforming viruses in three important properties: (1) they induce macroscopic neoplasms in infected animals only after a latent period of 4–12 months, compared with 2–3 weeks for the acute viruses; (2) they do not induce morphological transformation at detectable frequency in tissue culture cells; and (3) they do not seem to contain a transforming gene.

ALVs cause B-cell lymphomas and occasionally fibrosarcomas, nephroblastomas, erythroblastosis, osteopetrosis, and anaemia. Recent evidence suggests that these viruses cause neoplastic transformation by activating a normal cellular gene(s). (1) Many ALV-induced lymphomas do not contain viral 35S and/or 21S mRNAs^{1,2} and many of the integrated proviruses in these tumours are defective^{1–3}. Thus it seems unlikely that a viral gene product is required for maintenance of neoplastic transformation. (2) The oncogenic potential of ALV seems to reside within the 3' portion of the viral genome—within ~500 nucleotides from the poly (A) tract^{4–6}—a region which does not seem to encode a viral protein⁷. (3) ALV-induced lymphomas contain proviruses integrated at one or more sites. However, at least some proviral information in each tumour is found integrated at one of a limited number of common sites^{1–3}. As proviral integration is known to occur at a large number of sites on the host chromosome, possibly at random^{8–12}, this apparent specificity of integration in lymphomas suggests that neoplastic transformation requires integration adjacent to a specific gene(s). (4) Primary and metastatic tumours are clonal^{1–3,13}. Metastases in the same bird contain similar or identical patterns of proviral integration, based on restriction analyses of cellular DNA; thus, transformation is a rare event which generally occurs only once in each animal. (5) ALV-induced lymphomas and lymphoma-derived cell lines contain new tumour-specific RNAs consisting of viral 5'-terminal sequences covalently linked to cellular sequences^{1,2}. These RNAs fall into a limited number of size classes¹. (6) The tumour-specific RNAs are expressed at moderately high levels (100–300 copies per cell)¹.

Based on certain structural features of the integrated ALV provirus, we^{1,14} and others^{2,4,6,11} have suggested a model for oncogenesis consistent with all the above observations, which we have termed oncogenesis by promoter insertion (see Fig. 1). Studies in many laboratories have shown that the integrated provirus consists of the viral structural genes flanked by sequences of ~350 nucleotides, termed long terminal repeats (LTRs)^{15–18}. The LTR contains a putative promoter sequence^{7,12,19,20}. Initiation of normal viral RNA synthesis occurs within the left LTR. Because this sequence is repeated at the right end of the provirus, initiation could also occur within

the right LTR. If the provirus integrated adjacent to a potentially oncogenic cellular gene, transcription initiated from the viral promoter could generate an RNA molecule such as those found in ALV-induced lymphomas, containing both viral and cellular sequences (see Fig. 1). The resultant enhanced expression of this cellular gene would lead to neoplastic transformation. Identification of the putative cellular gene activated by this mechanism would be strong evidence in support of this model.

Over 10 different 'transforming genes' have been identified in the genomes of acute transforming retroviruses of avian and mammalian origin^{21–26}—each of these viral *onc* (*v-onc*) genes has a normal cellular counterpart (cellular *onc*, or *c-onc*)^{24,27–34}. These normal cellular genes are presumably the progenitors of the *v-onc* genes and are generally expressed at low levels in normal, uninfected cells^{29,35–38}, although they may be expressed at higher levels during certain stages of development³⁸. As they are highly conserved in vertebrate evolution^{25,29,39}, these normal cellular genes probably have some essential role in cellular growth, differentiation and/or development.

It seemed possible that the normal cellular gene(s) activated by promoter insertion might be one of the cellular counterparts of a *v-onc* gene. Here we report our efforts to identify the putative *c-onc* gene(s) involved in ALV-induced lymphomas, using cDNA probes specific for five *v-onc* genes of avian acute

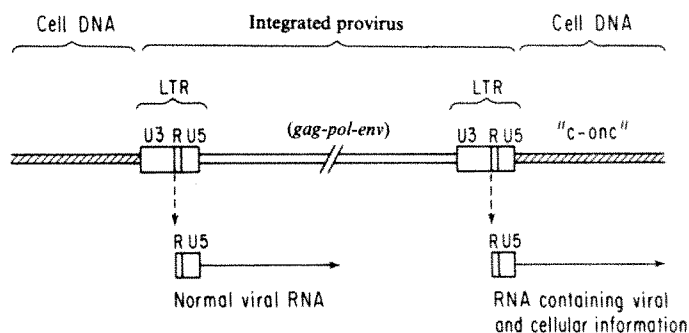


Fig. 1 Structure and transcriptional products of the integrated ALV provirus. The integrated provirus is flanked by sequences of ~350 nucleotides termed long terminal repeats (LTRs^{15–18}). The LTRs, shown greatly enlarged to emphasize structural details, are composed of sequences derived from the 5' (U5) and 3' (U3) ends of genomic RNA, and a short sequence (R) of ~20 nucleotides, present at both ends of genomic RNA. Synthesis of normal viral RNA (genomic RNA and mRNAs) initiates within the left LTR. Initiation within the right LTR would generate a molecule containing viral 5' sequences (R + U5) plus cellular information encoded in the adjacent cellular DNA. If, as shown, the provirus integrated upstream from a potentially oncogenic cellular gene (designated *c-onc*), initiation within the right LTR could cause elevated expression of the *c-onc* gene.

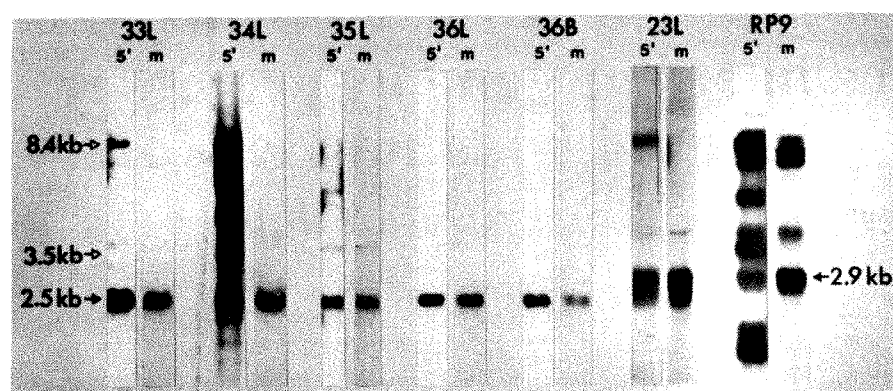


Fig. 2 *myc*-specific and viral 5'-specific RNAs in ALV-induced lymphomas. Poly(A)-containing RNAs (2–5 µg per lane) were glyoxylated, size-fractionated on 1% agarose gels, and transferred to aminophenylthioether paper by the procedure of Alwine *et al.*⁴³, as modified by B. Seed (personal communication). Details of this procedure have been described elsewhere¹. RNAs containing viral 5'-specific and *myc*-specific sequences were identified by hybridization to ³²P-labelled cDNA_{5'} (5') and cDNA_{myc} (m) (specific activity ~2–4 × 10⁸ c.p.m. per µg). The blot was first hybridized to cDNA_{5'} and exposed to X-ray film; it was then washed in denaturing conditions to remove the cDNA probe, and re-hybridized to cDNA_{myc}. Positions of viral 35S (8.4 kilobases) and 21S (3.5 kilobases) RNAs are indicated by open arrows; 2.5- and 2.9-kilobase *myc*-specific RNAs are indicated by closed arrows. Tissues are designated as in Tables 1 and 2. kb, Kilobase.

transforming viruses. We used three criteria to identify such a gene, based on predictions of the model for oncogenesis described above: (1) the *c-onc* gene should be expressed at elevated levels in ALV-induced lymphomas; (2) the *c-onc* sequences should be encoded in the tumour-specific RNAs, covalently linked to viral sequences; (3) proviruses in ALV-induced lymphomas should be integrated adjacent to the *c-onc* gene. Based on all three criteria, the data strongly suggest that most of the ALV-induced lymphomas studied result from activation of the *c-myc* gene, the cellular counterpart of the transforming gene of MC29 virus, by a promoter insertion mechanism.

Expression of *c-onc* genes in normal and neoplastic tissues

RNA from normal tissues and ALV-induced lymphomas was analysed by nucleic acid hybridization to cDNA probes specific for the following *v-onc* genes: *src* (Rous sarcoma virus), *erb* (avian erythroblastosis virus), *myb* (avian myeloblastosis virus), *fps* (Fujinami sarcoma virus) and *myc* (myelocytomatosis virus-29, or MC29). The cellular homologues of the five *onc* genes were expressed at low but detectable levels (1–4 copies per cell) in normal tissues from 3–4-month-old uninfected birds (Table 1). With one exception (see below), the cellular *src*, *erb*, *myb* and *fps* genes were expressed at similarly low levels in lymphomatous tissues.

In contrast, levels of *myc*-specific RNA were 30–100-fold higher in ALV-induced lymphomas (90–330 copies per cell). These data suggest that *c-myc* is the cellular gene responsible for ALV-induced neoplastic transformation in these tumours. An even higher level of *myc* RNA was found in RP9 (1,000 copies per cell), a cell line derived from an ALV-induced lymphoma. In RP9 cells, *myb*-specific RNA was also elevated, but to a much lesser extent (40 copies per cell); the significance of this is unclear.

A more extensive analysis of *myc* gene expression involving a larger number of normal tissues and ALV-induced lymphomas is shown in Table 2. Each tissue was also analysed for expression of virus-specific information by hybridization to a probe representative of the entire viral genome (cDNA_{rep}), and to a probe corresponding to the 5'-terminal 101 nucleotides of viral RNA (cDNA_{5'}); the latter would detect 5'-terminal (R+U5) sequences of RNAs initiated from either the left or the right LTR (see Fig. 1). As 5'-terminal sequences comprise only ~1% of the viral genome, the presence of viral 5' sequences alone would not be detected by cDNA_{rep}. For detection by cDNA_{rep}, an RNA molecule must contain other viral genetic information.

Six main points can be drawn from the data in Table 2: (1) 12 of 14 ALV-induced lymphomas (including the lymphoma cell line RP9) contained elevated levels of *myc* RNA (at least 80 copies per cell). One additional lymphoma (29L) contained a slightly elevated level (9 copies per cell), whereas sample 25L contained only normal levels of *myc* RNA (2 copies per cell). (2) Uninfected chicken embryo fibroblasts and bursal and muscle tissues from 3–6-month-old uninfected birds contained low levels of *myc*-specific RNA (2–5 copies per cell). (3) In infected,

but non-neoplastic tissues (5B, 6B, 19M, 23M), *myc*-specific RNA levels were similar to those in tissues from uninfected birds. Viral RNA is present at elevated levels in these tissues. Thus it is clear that viral infection and/or expression of viral genes alone is not sufficient to induce *myc* expression. (4) Viral gene expression is not necessary for maintenance of neoplastic transformation. In ~50% of the tumours, cDNA_{rep}-specific RNA was present at levels (one to six copies per cell) only slightly higher than those of the endogenous provirus *ev1*⁴⁰ in normal cells⁴¹. (Most uninfected chicken cells contain genetic information homologous to the replicative genes of RNA tumour viruses. These 'endogenous proviruses' (*ev*) segregate as defined genetic loci.) (5) In each of the samples containing low amounts of RNA homologous to cDNA_{rep} (7K, 10B, 23L, 25L, 35L, 36B, 36L), viral 5'-specific sequences were present at elevated levels. In all these tumours (except 25L, in which *myc* gene expression is not elevated) the level of 5'-specific RNA was similar to that of *myc*-specific RNA. This is consistent with the interpretation that 5' and *c-myc* sequences are present on the

Table 1 Transformation-specific RNAs in ALV-induced lymphomas

Sample no.	Tissue	RNA (copies per cell)				
		<i>src</i>	<i>fps</i>	<i>erb</i>	<i>myb</i>	<i>myc</i>
12B	Uninfected	3	1	2	3	3
31L	Uninfected	2	1	4	1	2
7K	Lymphoma	4	1	6	5	110
11B	Lymphoma	3	2	2	4	180
23L	Lymphoma	5	1	2	3	330
9B	Lymphoma	5	1	4	2	230
22L	Lymphoma	4	2	ND	5	90
RP9	Lymphoma cell line	3	1	3	40	1000

Sample designations indicate the animal number and the source of the tissue: B, bursa; L, liver; K, kidney. RP9 is a continuous cell line derived from a RAV-2-induced lymphoma (provided by Dr W. Okazaki). Lymphomas were removed from leukotic birds 4–6 months after infection with ALV. Birds were injected with 10⁶–10⁷ plaque-forming units (PFU) of virus, 2–7 days after hatching. Virus strains used in this and other experiments included Rous-associated virus-1 (RAV-1), RAV-2 and transformation-defective (td) mutants of Schmidt-Ruppin Rous sarcoma virus-A (td107A, td103 and td109). No virus strain-specific differences were observed in any of the experiments described here. Normal tissues were taken from uninfected 3–6-month-old birds. Hybridizations were performed in liquid as described elsewhere³⁵, using ~5 mg ml⁻¹ of cellular RNA. Hybrid reactions were stopped at various time intervals, corresponding to C_t/t values of ~10–10⁵ mol·s⁻¹. RNA concentrations were calculated from C_t/t values (see ref. 35), assuming that a C_t/t value of 4 × 10⁴ mol·s⁻¹ is equivalent to one copy per cell. Because of variations in the RNA content of cells from different tissues, the values shown are not precise but the error introduced is no more than a factor of 2–3, and thus does not affect any of the conclusions drawn. Radiolabelled cDNA probes correspond to the unique (putative transforming) sequences of the following acute transforming viruses^{24,27,61,62}: Rous sarcoma virus (*src* gene), Fujinami sarcoma virus (*fps*), avian erythroblastosis virus (*erb*), avian myeloblastosis virus (*myb*), and myelocytomatosis virus-29 (*myc*). Probes were prepared as previously described for cDNA_{rep} (ref. 35), except that cDNA was synthesized from purified viral RNA and reverse transcriptase, using random DNA primer⁶³. Sequences unique to each of the acute transforming viruses were selected by removing sequences common to ALV (by hybridization to a mixture of ALV RNAs from RAV-1, RAV-2, RAV-61, plus erythroblastosis-associated virus for cDNA_{erb}, and myeloblastosis-associated viruses-1 and -2 for cDNA_{myb}). The probes demonstrated no cross-homology, and did not hybridize (<2%) with any of the ALV RNAs used for their selection, or to ribosomal RNA. ND, not determined.

same RNA molecules in ALV-induced lymphomas. (6) *c-myc*-specific RNA in normal cells does not contain viral 5' sequences—the latter were undetectable (less than 0.1 copy per cell) in tissues from an ev-negative bird (37L and 37B)⁴², whereas *myc*-specific RNA was present at 2 and 5 copies per cell, respectively.

Tumour-specific RNAs have *myc*-specific sequences and viral 5' sequences

To test directly whether the *myc*-specific sequences and viral 5'-sequences are encoded on the same RNA transcript, poly(A)-containing RNA from ALV-induced lymphomas was analysed by the 'Northern' transfer technique⁴³, using cDNA_{5'} and cDNA_{myc} as probes. Previous studies showed that almost all tumours synthesize discrete new tumour-specific RNAs containing viral 5' sequences plus other information, presumably derived from the cell¹. In most tumours, the size of these RNAs was ~2.5 kilobases, although a few contained RNA of ~2.9 kilobases. None of these RNAs contained sequences detectable with cDNA_{rep}. In a few tumours, larger RNAs of 5–6 kilobases were detected with cDNA_{5'}. These RNAs also contained information detectable with cDNA_{rep}.

Six of the seven samples shown in Fig. 2 contained a single major *myc*-specific RNA. In each case, this RNA corresponded to a species detectable with cDNA_{5'}, suggesting that the *myc* sequences and viral 5' sequences are present on the same RNA transcript. As reported previously¹, these RNAs are not detected with cDNA_{rep}. In five tumours, the size of this RNA was 2.5 kilobases, whereas in a sixth (23L), it was rather heterogeneous, with a size of ~2.9 kilobases. A 2.9 kilobase RNA was also found in RP9 cells, which also contained four additional *myc*-specific RNA species, each of which corresponded to a 5'-containing RNA. The three larger of these, with sizes of 7–8.5 kilobases, were rather obscured in the cDNA_{5'}-specific analyses by the presence of normal viral 35S RNA, but were seen clearly in short exposures of the blots. The high degree of specificity of the cDNA_{myc} probe can be seen from the fact that this probe did not hybridize to the 35S viral RNA in tumour 34L, although this species is much more abundant than the 2.5-kilobase RNA detected with cDNA_{5'} and cDNA_{myc}. RP9 and 34L also contain several smaller RNAs that hybridize with cDNA_{5'} and cDNA_{rep} (ref. 1) but not with cDNA_{myc}; these RNAs are presumably aberrant viral transcripts.

myc-specific RNA was not detected in normal cells under the conditions used in these experiments (data not shown). Sheiness *et al.*⁴⁴ have identified *c-myc* RNA in uninfected cells using larger amounts of cellular RNA. These workers reported a size of 2.8 kilobases, slightly larger than that of the most frequently observed *myc*-specific RNA found in ALV-induced lymphomas.

Genetic content of tumour-specific RNAs

The data described above are consistent with the interpretation that the *myc*-specific RNAs in ALV-induced lymphomas have the structure illustrated in Fig. 1 (that is, they contain only ~100 nucleotides of virus-coded information). The failure to detect these RNAs with cDNA_{rep} (ref. 1) suggests that they contain little or no viral information other than the 5'-terminal sequences detected with cDNA_{5'}. As analysis with cDNA_{rep} does not provide a sensitive assay for possible short stretches of viral information, RNA from ALV lymphomas was tested further by liquid hybridization, using probes specific for portions of the viral genome.

The low level of normal viral mRNAs in some of the tumours allowed us to examine the genetic content of the *myc*-specific RNA in these tissues. Two tumours, 7K and 23L, in which *myc*-specific RNA was at least 50-fold more abundant than viral mRNA (Table 2), were chosen. RNA from these tumours was hybridized to cDNA_{myc} (Fig. 3, broken lines) and to probes specific for the viral *gag*, *pol*, and *env* genes, the 500 nucleotides adjacent to the poly(A) tract at the 3' end of the viral genome

(cDNA_{3'}), and cDNA_{5'} (solid lines). With both tumours, cDNA_{myc} hybridized to a plateau level of ~85%, indicating extensive homology between the *myc* probe and the *c-myc* sequences encoded in the tumour-specific RNA of ALV-induced lymphomas. A plateau value of ~90% hybridization was obtained with RNA from normal cells, when hybridization was carried out to very high *C_t* values (data not shown).

Of the virus-specific probes, only cDNA_{5'} hybridized with kinetics similar to those obtained with cDNA_{myc}. Some hybridization was obtained with the other probes at much higher *C_t* values. This could reflect the presence of small amounts of viral mRNA. However, in no case were the hybridization patterns consistent with the presence of any viral sequences, other than 5'-specific sequences, in the *myc*-specific RNA transcript. Thus we conclude that, within the limits of this assay (probably ~5% of the genetic content of each probe), no viral information other than 5' sequences is present in the *myc*-specific transcripts of tumours 7K and 23L. Hybridizations of gene-specific probes to the 'Northern' blots similarly failed to detect other viral sequences on the 2.5- and 2.9-kilobase *myc*-specific RNAs (data not shown).

ALV provirus is integrated adjacent to *c-myc* in ALV-induced lymphomas

In our previous studies¹, analysis of lymphoma DNA was performed by digestion with the restriction endonuclease *EcoRI*, which cuts at several sites within the provirus^{15,5,45}. The most useful restriction fragment generated by this digestion is the right-end 'junction fragment', which contains ~150 nucleotides of viral information, plus adjacent downstream cellular DNA. If the ALV provirus is integrated adjacent to *c-myc*, this junction

Table 2 *myc*-specific and virus-specific RNA in normal and neoplastic tissues

Sample no.	Tissue	RNA (copies per cell)		
		<i>myc</i>		<i>rep</i>
2928F	Normal (uninfected)	2	18	0.7
31L	Normal (uninfected)	2	6	0.5
12B	Normal (uninfected)	3	07	0.7
37L	Normal (uninfected)	2	<0.1*	<0.1*
37B	Normal (uninfected)	5	<0.*	<0.1*
19M	Normal (infected)	2	1,500	1,300
23M	Normal (infected)	3	1,500	1,200
5B	Normal (infected)	3	145	160
6B	Normal (infected)	2	200	280
3B	Pre-neoplastic bursa	10	7	0.4
7K	Lymphoma	110	100	2
9B	Lymphoma	230	1,300	1,100†
10B	Lymphoma	100	90	2
11B	Lymphoma	180	2,500	2,000
19L	Lymphoma	85	1,900	1,300
22L	Lymphoma	90	300	300
23L	Lymphoma	330	300	5
25L	Lymphoma	2	90	2
29L	Lymphoma	9	1,700	1,300
34L	Lymphoma	105	1,000	800
35L	Lymphoma	80	80	6
36B	Lymphoma	100	100	2
36L	Lymphoma	140	120	3
RP9	Lymphoma cell line	1,000	4,000	600

Sample designations are as in Table 1; F, fibroblasts (chicken embryo); L, liver; B, bursa; M, muscle; K, kidney. Tissues were taken from 3–6-month-old uninfected birds, or from leukotic infected chickens 4–6 months after infection (see Table 1 for details). RNA concentrations were determined as in Table 1. RNA was detected with cDNA_{myc}, or with virus-specific probes corresponding to 5'-terminal sequences (R+U5) of genomic RNA (cDNA_{5'}) or to the entire viral genome (cDNA_{rep}). cDNA_{5'} ('strong-stop DNA'⁶⁴) was synthesized from RAV2 as described elsewhere¹; cDNA_{rep} was synthesized from RAV-2 RNA, using random DNA primer⁶³. cDNA_{5'} detects RNA transcripts initiating either from the left or right LTR (see Fig. 1), whereas cDNA_{rep} detects only viral RNA containing substantial portions of the viral coding sequences (presumably initiating from the left LTR).

* No virus-specific RNA was detected with either cDNA_{5'} or cDNA_{rep} in conditions that would have detected 0.1 copies per cell. These samples were taken from a chicken (no. 37) which lacks endogenous proviruses⁴².

† RNA from tumour 9B hybridized with kinetics suggesting two components, with levels of 25 and 1,100 copies per cell. Other experiments demonstrated the presence in this tumour of an abundant 5.5-kilobase RNA species containing *gag* and 3' viral sequences, and portions of *pol*.

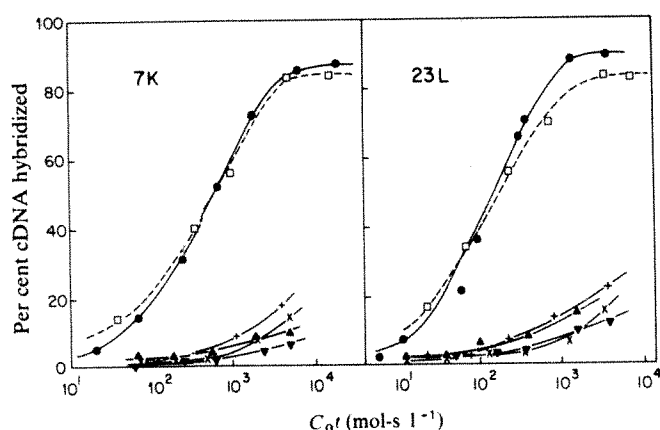


Fig. 3 Genetic content of *myc*-specific RNAs in two ALV-induced lymphomas. Total cellular RNA from tumours 7K and 23L was extracted with proteinase K-phenol-chloroform³⁵, and precipitated two or three times with 2 M LiCl to remove cellular DNA. RNA was hybridized to cDNA_{myc} (broken line) and to cDNA probes specific for different portions of the ALV genome (solid lines). The following virus-specific probes were used: ▲, cDNA_{gag}; ▼, cDNA_{pol}; ×, cDNA_{env}; ●, cDNA_{5'}; and +, cDNA_{3'}. cDNA_{pol} was prepared from RAV-0, and corresponds to the deletion in ev3 RNA⁴¹, which includes a large portion of the *pol* gene, and the 3' portion of *gag*. All other probes were prepared from RAV-2. cDNA_{gag} represents the 5' 2,500 nucleotides of genomic RNA, and overlaps somewhat with cDNA_{pol}. cDNA_{env} corresponds to the *env* deletion of Bryan strain of Rous sarcoma virus, and cDNA_{3'} corresponds to the 3' 500–600 nucleotides ('common' or 'c' region) of genomic RNA. cDNA_{5'} is described in Table 2 legend. Probes were prepared and hybridizations performed as described elsewhere^{35,41}.

fragment should contain *myc*-specific sequences, and thus be detected with both cDNA_{5'} and cDNA_{myc}.

Restriction analyses of DNA from normal tissue and from three ALV-induced lymphomas are shown in Fig. 4. In normal cells, *c-myc* sequences were contained almost entirely within a single *EcoRI* restriction fragment of ~14 kilobases, consistent with the report of Sheiness *et al.*⁴⁴. (We also found a second, minor band at ~2 kilobases which has not been further characterized.)

Lymphoma DNAs contained a new *myc*-specific restriction fragment not present in normal cells (Fig. 4). In each case, this new fragment corresponded to a junction fragment detected with cDNA_{5'}. This presumably represents integration of an ALV provirus adjacent to the *c-myc* gene. In a larger sampling of lymphomas, we found that 31 out of 37 lymphoma DNAs contained proviral information integrated adjacent to *c-myc*. The 14- (and 2-) kilobase *myc*-specific *EcoRI* fragments found in normal cells were also found in the lymphomas. This would be expected because integration presumably occurs adjacent to only one of the copies of *c-myc* present in the diploid chicken genome.

We have recently succeeded in cloning the *EcoRI* junction fragments from two ALV-induced lymphomas (B.G.N. *et al.*, unpublished results). The cloned DNAs contain both viral 5' information and *myc*-specific information, thus providing direct evidence that the viral and *c-myc* sequences are covalently linked in these tumours.

Discussion

From the above data we conclude that, in almost all ALV-induced lymphomas, neoplastic transformation results from activation of the cellular gene *c-myc*. This process seems to involve integration of the provirus upstream from the *c-myc* gene, and transcriptional activation of this gene due to the insertion of the 'strong' viral promoter. The resultant RNA transcript contains both viral and *c-myc* information, and is present at levels 30–100-fold higher than that of *c-myc* RNA in normal tissues.

Integration of the provirus at a site which could activate the *c-myc* gene would presumably be a rare event, because integra-

tion is known to occur at many sites (possibly at random) in the host genome^{8–12}. Thus, this mechanism can explain, at least in part, the long latent period required for development of lymphomas in ALV-infected birds, and the failure to detect transformation in ALV-infected tissue culture cells.

We have previously observed that ALV-induced lymphomas contained proviruses integrated at a limited number of discrete sites, as judged by restriction analysis. The junction fragments of different size observed in the earlier experiments could represent ALV integrations adjacent to different cellular genes. However, our data (see Fig. 4 and ref. 46) indicate that most proviruses are integrated adjacent to the same gene, *c-myc*. Apparently, integration at several locations upstream from *c-myc* can lead to activation of this gene. Many tumours that have junction fragments of different sizes (for example, 34L and 35L) synthesize *myc*-specific RNAs of the same size (2.5 kilobases). This could be explained if the 2.5-kilobase RNAs arose by splicing events, with removal of varying lengths of intervening sequences.

It is unclear why *c-myc* should be involved in such a large percentage (~85%) of lymphomas. Over 10 transforming genes (*v-onc*), each of which has a cellular counterpart that is highly conserved in all vertebrates, have been identified in different avian and mammalian retroviruses^{21–34}—others undoubtedly exist. Although some *v-onc* genes may show tissue specificity, this alone cannot explain the preferential involvement of *c-myc* in B-cell lymphomas. MC29, an acute transforming virus which contains the *myc* gene, causes myelocytomatosis, fibrosarcomas and carcinomas in infected birds⁴⁷. No examples of B-cell lymphomas have been reported. In tissue culture, MC29 transforms both haematopoietic cells⁴⁸ and fibroblasts^{47,48}. However, other viruses (for example, AMV and AEV) also transform haematopoietic cells *in vivo* and *in vitro*²³. Nevertheless, the cellular counterparts of these *v-onc* genes are not activated by ALV in a large percentage of ALV-induced lymphomas.

Breindl *et al.*⁴⁹ have suggested that proviruses integrate preferentially into transcriptionally active regions of chromosomes. If *c-myc* is in a transcriptionally active conformation at some stage of B-cell differentiation, integration within this region might occur with a higher than random frequency. Alternatively, *c-myc* may be more readily activated by the promoter insertion mechanism, due to some structural feature(s) of this

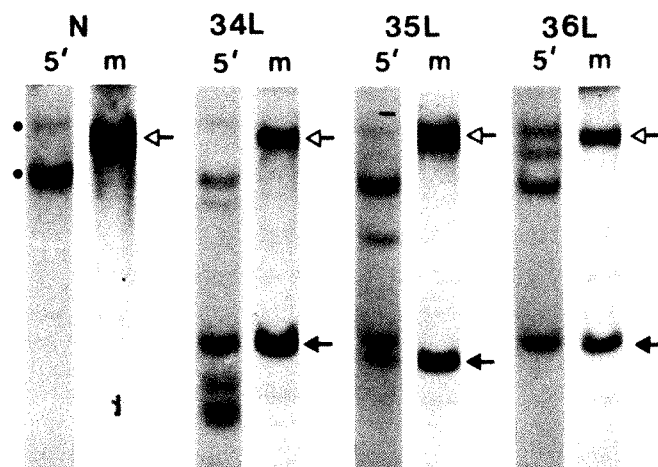


Fig. 4 Restriction analysis of *myc*-specific and viral 5'-specific sequences in DNA from normal cells and ALV-induced lymphomas. DNA was extracted from normal brain (N) from bird no. 34, and from three ALV-induced lymphomas (designated as in Tables 1 and 2). DNA was digested with restriction endonuclease *EcoRI*, size-fractionated on agarose gels and transferred to nitrocellulose paper by the technique of Southern⁶⁵. Experimental details have been described elsewhere^{1,40}. Duplicate samples from each tissue were analysed in parallel experiments by hybridization to cDNA_{5'} (5') or cDNA_{myc} (m). Two *myc*-specific restriction fragments—a major species of 14 kilobases (open arrows) and a minor species of 2 kilobases—were detected in all DNAs from both normal and neoplastic tissues. *myc*-specific fragments of 2.9 and 2.7 kilobases (solid arrows), also detected with cDNA_{5'}, were found only in tumour DNAs. cDNA_{5'}-specific restriction fragments of the endogenous provirus ev1^{1,40} (indicated in the left lane by solid circles) were present in both normal and transformed tissues.

gene. As discussed above, integration at several sites upstream from *c-myc* can apparently result in its activation. This multiplicity of potentially 'productive' sites might increase the probability of activating this gene over other *c-onc* genes, which might only be activated by integration at a more limited number of sites.

It is possible that other *c-onc* genes are activated by promoter insertion in the ALV-induced lymphomas that do not show enhanced *myc* expression. We are currently attempting to identify the cellular gene(s) induced in these tumours.

We could find no evidence for the participation or generation of an acute transforming virus in ALV-induced lymphomas; no virus capable of transforming fibroblasts was recovered from any of the lymphomas tested (data not shown). (MC29 transforms fibroblasts in tissue culture with high efficiency.) Virus from one sample, RP9, was further tested by injection into 1–7-day-old chicks. Lymphomas appeared after a long latent period (>4 months), but no acutely developing neoplasia was observed. In addition, the structure of the *myc*-specific RNAs in lymphomas was not consistent with that of a viral genome. All known defective transforming viruses contain viral sequences at the 5' and 3' ends of their genome. Sequences at the 3' end encode recognition signals for reverse transcription of viral RNA, integration of proviral DNA and synthesis of viral mRNA. Analyses of RNA from two ALV-induced lymphomas (Fig. 3) revealed no viral 3' information. Furthermore, sequences thought to be involved in RNA packaging⁵⁰, located in the region from ~100–350 nucleotides from the 5' end⁵¹, do not seem to be present in the tumour-specific RNAs.

Cooper and Neiman⁵² using a transfection assay, recently demonstrated that DNA from ALV-induced lymphomas causes transformation of NIH/3T3 cells with high efficiency. Surprisingly, the transformants did not contain any detectable viral DNA sequences (including the LTR), nor apparently did they contain *myc*-specific sequences derived from the lymphoma DNAs. However, these investigators have recently found that the lymphomas from which the DNA was derived do contain ALV proviruses integrated adjacent to *c-myc* (G. Cooper and P. Neiman, personal communication). The transfection assay is apparently detecting another genetic change in lymphomas which may be a direct or indirect consequence of enhanced *c-myc* expression. It is possible that *c-myc* activation, and the activity detected in the transfection assay, represent different steps in a multistage process leading to neoplastic transformation.

The *onc* gene of MC29 virus^{53–55}, and those of many other acute transforming viruses⁵⁶, are expressed as *gag-onc* fusion proteins. Our data indicate that the *gag*-encoded portion of this protein is not required for the transforming activity of the *myc* gene product at least in B cells. We could detect no viral sequences other than the 5'-terminal 100 nucleotides in *myc*-specific RNA from these tumours; thus the predicted translation product of this RNA would lack *gag* sequences.

Earlier work had suggested that normal cellular information might be potentially oncogenic. Hanafusa *et al.*⁵⁷ generated acute transforming viruses (recovered avian sarcoma viruses, rASVs), by injecting transformation-defective viruses which lack most, but not all, of the *src* gene, into chickens. The rASVs have acquired most of their *src* information from the cell by recombination with *c-src*^{58–60}. These experiments suggested that *c-src*, without substantial modification, is capable of inducing cell transformation when expressed at high levels as a consequence of its incorporation into the viral genome. Recently, Oskarsson *et al.*³³ have shown that *c-mos*, the cellular counterpart of the transforming gene of Moloney murine sarcoma virus, can induce transformation of NIH/3T3 cells if the LTR is linked to cloned *c-mos* DNA. The resultant structure is similar to that found in ALV-induced lymphomas. No viral sequences other than the LTR are required for this activity (G. Vande Woude, personal communication).

It seems likely, therefore, that the role of both acute and slow RNA tumour viruses in neoplastic transformation is simply to provide a means for constitutively expressing a cellular gene that is normally expressed at low levels. These viruses accomplish this in slightly different ways. In the acute viruses, a cellular gene has been inserted into the viral genome (by recombination), where it comes under the control of the viral promoter. With slow transforming viruses such as ALV, the viral promoter is inserted into (or adjacent to) the cellular gene.

The data presented here show, for the first time, that a carcinogenic agent (ALV) acts by altering the expression of a normal cellular gene. These findings raise the interesting possibility that activation of a *c-onc* gene may be the common initiation event in neoplastic transformation by both viral and non-viral agents. Radiation and nearly all chemical carcinogens are potent mutagens—these might induce elevated expression of a *c-onc* gene by altering regulatory sequences (by point mutation or DNA modification), and/or by causing translocations that join *c-onc* coding sequences to transcriptionally active elements. The different types of cancer may reflect the particular *c-onc* gene activated, and the type of cell in which activation occurs. The availability of molecular probes for the known *onc* genes should allow screening for enhanced expression of these genes in a large number of tumours of unknown origin. Thus it should be possible to test the general validity of this model.

We thank N. Goldberg and J. Fang for technical assistance; M. Shibuya for providing *fps* probe; J.-H. Chen and S. M. Anderson for providing virus strains and RNAs; H. Hanafusa, T. Hanafusa, and L.-H. Wang for supplying some of the leukotic birds; W. Okazaki for the cell line RP9, J. Walberg for histopathological analyses; H. Hanafusa and J. Summers for critical reading of the manuscript, and G. Gasic for helpful discussions.

This work was supported by grants from the NIH and the ACS. B.G.N. is a biomedical fellow in the MSTP of the NIH.

Received 29 December 1980; accepted 9 February 1981.

- Neel, B. G., Hayward, W. S., Robinson, H. L., Fang, J. M. & Astrin, S. M. *Cell* **23**, 323–334 (1981).
- Payne, G. S. *et al. Cell* **23**, 311–323 (1981).
- Fung, Y. K., Fadly, A. M., Crittenden, L. B. & Kung, H.-J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Robinson, H. L., Pearson, M. N., DeSimone, D. W., Tschlis, P. N. & Coffin, J. M. *Cold Spring Harb. Symp. quant. Biol.* **44**, 1133–1142 (1980).
- Crittenden, L. B., Hayward, W. S., Hanafusa, H. & Fadly, A. M. *J. Virol.* **33**, 915–919 (1980).
- Tschlis, P. N. & Coffin, J. M. *J. Virol.* **33**, 238–249 (1980).
- Czernilovsky, A. P. *et al. Nucleic Acids Res.* **13**, 2967–2984 (1980).
- Hughes, S. H. *et al. Cell* **15**, 1387–1410 (1978).
- Steffen, D. & Weinberg, R. A. *Cell* **15**, 1003–1010 (1978).
- Ringold, G. M., Shank, P. R., Varmus, H. E., Ring, J. & Yamamoto, K. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 665–669 (1979).
- Quintrell, N., Hughes, S. H., Varmus, H. E. & Bishop, J. M. *J. molec. Biol.* **143**, 363–393 (1980).
- Shimotohno, K., Mizutani, S. & Temin, H. M. *Nature* **185**, 653–654 (1980).
- Neiman, P., Payne, L. N. & Weiss, R. A. *J. Virol.* **34**, 178–186 (1980).
- Hayward, W. S., Neel, B. G., Fang, J., Robinson, H. L. & Astrin, S. M. *Mod. Trends in Tumour Leukemia* (1980).

- Hsu, J. W., Sabran, J. L., Mark, G. E., Guntaka, R. V. & Taylor, J. M. *J. Virol.* **28**, 810–818 (1978).
- Shank, P. R. *et al. Cell* **15**, 1383–1395 (1978).
- Sabran, J. L. *et al. J. Virol.* **29**, 170–178 (1979).
- Hughes, S. H. *et al. Cell* **18**, 347–359 (1979).
- Yamamoto, T., Jay, G. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 176–180 (1980).
- Ju, G. & Skalka, A. M. *Cell* **22**, 379–386 (1980).
- Hanafusa, H. *Comprehensive Virol.* **10**, 401–483 (1977).
- Vogt, P. K. *Comprehensive Virol.* **9**, 341–455 (1977).
- Graf, T. & Beug, H. *Biochem. biophys. Acta* **516**, 269–299 (1978).
- Roussel, M. *et al. Nature* **281**, 452–455 (1979).
- Fischinger, P. J. in *Molecular Biology of RNA Tumor Viruses* (ed. Stephenson, J. R.) 3705–3709 (Academic, New York, 1980).
- Shih, T. Y. & Scolnick, E. M. in *Viral Oncology* (ed. Klein, G.) 135–160 (Raven, New York, 1980).
- Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. *Nature* **260**, 170–173 (1976).
- Frankel, A. E. & Fischinger, P. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3705–3709 (1976).
- Sheiness, D. K. & Bishop, J. M. *J. Virol.* **31**, 514–521 (1979).
- Chien, Y.-H. *et al. J. Virol.* **31**, 752–760 (1980).
- Baltimore, D. *et al. Cold Spring Harb. Symp. quant. Biol.* **44**, 849–854 (1980).
- Simek, S. & Rice, N. J. *J. Virol.* **33**, 320–329 (1980).
- Oskarsson, M., McClements, W. L., Blair, D. G., Maizel, J. V. & Vande Woude, G. F. *Science* **207**, 1222–1224 (1980).

34. Yoshida, M., Kawai, S. & Toyoshima, K. *Nature* **287**, 653–654 (1980).
35. Hayward, W. S. *J. Virol.* **25**, 47–63 (1977).
36. Wang, S. Y., Hayward, W. S. & Hanafusa, H. *J. Virol.* **24**, 64–73 (1977).
37. Spector, D. H. *et al. Cell* **13**, 371–379 (1978).
38. Chen, J.-H. *J. Virol.* **36**, 162–170 (1980).
39. Spector, D. H., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4102–4106 (1978).
40. Astrin, S. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5941–5945 (1978).
41. Hayward, W. S., Braverman, S. B. & Astrin, S. M. *Cold Spring Harb. Symp. quant. Biol.* **44**, 1111–1121 (1980).
42. Astrin, S. N., Buss, E. G. & Hayward, W. S. *Nature* **282**, 339–341 (1979).
43. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350–5354 (1977).
44. Sheiness, D. K., Hughes, S. H., Varmus, H. E., Stubblefield, E. & Bishop, J. M. *Virology* **105**, 415–424 (1980).
45. Boone, L. R. & Skalka, A. M. *J. Virol.* (in the press).
46. Astrin, S. M., Neel, B. G., Rogler, C., Skalka, A. M. & Hayward, W. S. (in preparation).
47. Langlois, A. J., Beard, D. & Beard, J. W. in *Comparative Leukemia Research* (ed. Dutcher, R. M.) 96–105 (Karger, Basle, 1970).
48. Graf, T. *Virology* **54**, 398–413 (1973).
49. Breindl, M., Bachelier, L., Fan, H. & Jaenisch, R. *J. Virol.* **34**, 373–382 (1980).
50. Linial, M., Meideros, E. & Hayward, W. S. *Cell* **15**, 1371–1381 (1978).
51. Shank, P. & Linial, M. *J. Virol.* **36**, 450–456 (1980).
52. Cooper, G. M. & Neiman, P. E. *Nature* **287**, 656–658 (1980).
53. Bister, K., Hayman, M. J. & Vogt, P. K. *Virology* **82**, 431–448 (1977).
54. Kitchener, G. & Hayman, M. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1637–1641 (1980).
55. Mellon, P., Pawson, A., Bister, K., Martin, G. S. & Duesberg, P. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5874–5878 (1978).
56. Stephenson, J. R. in *Molecular Biology of RNA Tumor Viruses* (ed. Stephenson, J. R.) 245–297 (Academic, New York, 1980).
57. Hanafusa, H., Halpern, C. C., Buchhagen, D. L. & Kawai, S. *J. exp. Med.* **146**, 1735–1747 (1977).
58. Wang, L.-H., Halpern, C. C., Nadel, M. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5812–5816 (1978).
59. Karess, R. E., Hayward, W. S. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3154–3158 (1979).
60. Karess, R. E., Hayward, W. S. & Hanafusa, H. *Cold Spring Harb. Symp. quant. Biol.* **44**, 765–771 (1980).
61. Hanafusa, T., Wang, L.-H., Anderson, S. M., Karess, R. E. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3009–3013 (1980).
62. Lee, W.-H. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2018–2022 (1980).
63. Taylor, J. M., Illmensee, R. & Summers, J. *Biochim. biophys. Acta* **442**, 324–330 (1976).
64. Haseltine, W. A., Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 989–993 (1977).
65. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).

LETTERS

A companion quasar to 3C345

Bruce Margon*, Gary A. Chanan†
& Ronald A. Downes‡

*Department of Astronomy, University of Washington, Seattle, Washington 98195, USA

†Columbia Astrophysics Laboratory, Columbia University, New York, New York 10027, USA

‡Department of Astronomy, University of California, Los Angeles, California 90024, USA

Although some quasistellar objects (QSOs) are known to be members of small groups of galaxies (see refs 1–3), there is no known case of a QSO in a rich cluster of galaxies⁴. Whether this has important implications for the nature of QSOs, or is simply a reflection of the difficulty of observing high-redshift clusters, is unclear. Here we report the discovery of a new QSO which provides perhaps the strongest evidence to date for quasar membership in a rich cluster.

The QSO was discovered during a programme aimed at the identification of optical counterparts of X-ray sources observed serendipitously by the Einstein Observatory. The detector used, the imaging proportional counter⁵, yields positional accuracies of the order of 1 arc min, typically restricting optical candidates for the X-ray source to only a few objects. These candidates were then each examined spectroscopically. So far we have identified more than 40 previously uncatalogued QSOs using this technique, and preliminary results on the objects as a class are reported elsewhere⁶. During X-ray observations of quasar 3C345, a serendipitous X-ray source located 8 arc min north-east of the quasar was discovered, in addition to emission from 3C345 itself. The observed flux from the new source is 3.2×10^{-13} erg cm⁻² s⁻¹ in the 0.5–4.5-keV band. This flux, which is roughly 7% of that observed from 3C345 (ref. 7), is typical of the serendipitous sources in our sample. Three separate observations during a 6-month period in 1979–80 showed no evidence for intensity variability of the new source in excess of the $\pm 20\%$ statistical uncertainty in the count rate. There was also no evidence for X-ray variability in 3C345 during this period.

Spectroscopic observations of optical candidates in the X-ray field were obtained on 16 May 1980, using the Robinson-Wampler image tube scanner⁸ at the Lick Observatory 3-m Shane reflector; the grating yields 20 Å spectral resolution in the 3,500–8,500 Å range. The X-ray source was readily identified with a previously unreported faint QSO, which we designate 1641+3998. The coordinates of the object are $\alpha(1950) = 16$ h 41 min 54.1 s, $\delta = +39^\circ 59' 17''$. Based on calibrations of the image diameter of the object on the Palomar Observatory Sky Survey⁹, we estimate the magnitude of the object at the epoch of

those plates as $B = 18.7 \pm 0.5$. Both the Sky Survey and our spectrophotometry show the QSO to be quite blue. As stressed in ref. 6, the probability of a chance coincidence of a random QSO with the X-ray source is negligible, so the identification of the quasar with the source is secure. The finding chart for the object in Fig. 1 is reproduced from a 4-m Mayall telescope plate obtained in 1973; also shown is 3C345.

The spectrum of the newly discovered QSO is of particular interest, and is shown in Fig. 2. Despite the modest signal-to-noise ratio, four strong emission lines are clearly visible at $\lambda\lambda$ 4,450, 6,912, 7,750, and 7,980 Å. We identify these features as Mg II λ 2,800, H γ , H β and [O III] λ 5,007, respectively, and derive a galactocentric redshift $z = 0.594 \pm 0.001$, where the uncertainty is dominated by the breadths of the permitted emission lines. The galactocentric redshift of 3C345 (see refs 10, 11) is $z = 0.595$, where again we estimate a formal uncertainty of at least ± 0.001 due to emission line widths. Thus to within the measurement errors, the redshifts of these two quasars are identical. The spectra of the objects are at least superficially similar in appearance, although this is perhaps not surprising in view of the limited variety in the emission lines which are typically seen in QSOs at this redshift.

We may estimate the probability that the remarkable redshift agreement of the two objects is a chance coincidence of two randomly distributed, unrelated QSOs. Our X-ray observations⁶ yield 0.6 serendipitous sources per square degree, and we identify about half of these as QSOs with $z \leq 1$. Thus we expect to observe $\approx 6 \times 10^4$ QSOs before finding one with a

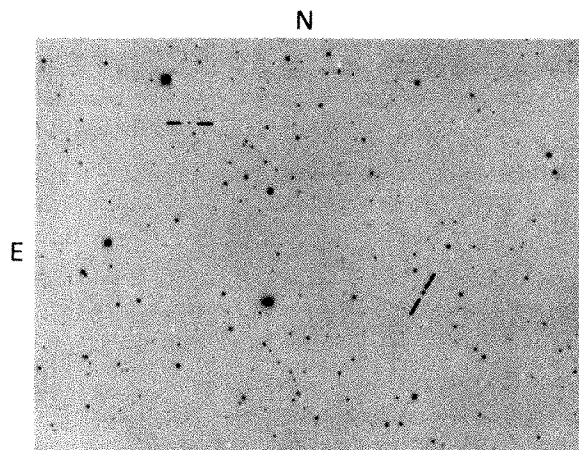


Fig. 1 A blue plate (IIa0 + GG385) of the field of the two quasars, obtained by Kinman with the Mayall 4-m telescope on 3 June 1973. The newly discovered quasar, 1641+3998, is the fainter of the two flagged objects, near the north-east of the field; 3C345 is the brighter object marked 8 arc min to the south-west.

'companion' as close in position (0.06 deg^2) and redshift (10^{-3}) as the current example is to 3C345. In fact only about 100 QSOs had been observed⁷ when 1641+3998 was discovered, so the probability of a chance coincidence is ~ 0.002 . Therefore, tempered by the hazards of a *posteriori* probability calculations, we conclude that the association of 3C345 and 1641+3998 is statistically significant and unlikely to be accidental.

If we accept the physical association of our two quasars, the explanations for such a situation are limited. We have briefly considered the extraordinary possibility that 1641+3998 could be a gravitationally focused image of 3C345, similar to the two currently known examples of this effect^{12,13}. The very large angular separation of the images compared with the previous cases would imply an unusual mass and/or mass distribution for the focusing matter. Remarkably this situation cannot be immediately dismissed, as a cluster of galaxies at $z=0.01$ with mass $2 \times 10^{14} M_{\odot}$ could create such an image. However, such a cluster must have a very extreme mass-to-luminosity ratio to avoid obvious detection. A more awkward difficulty with this hypothesis is that the ratio of fluxes of the two QSOs should be independent of wavelength. Although the optical and X-ray flux ratios of the two objects do agree well (within 3 orders of magnitude), 3C345 is an intense radio source (11 Jy at 178 MHz) leading to the prediction of an easily detectable radio flux from 1641+3998, barring long differential time delays. Maps which subtract the side lobes of 3C345 (ref. 14) do not show such a source, to limits of 10 mJy. The agreement of the ratio of optical and X-ray fluxes is thus probably simply a manifestation of the relative constancy of this parameter previously noted in large samples of radio-selected¹⁵ as well as X-ray selected⁶ QSOs.

The alternative hypothesis, which we favour, is that the redshift agreement of these two QSOs indicates their common membership in a cluster of galaxies. The projected separation of several Mpc would then indicate this to be a rich cluster, the first such example of a cluster with QSOs. Although theoretical calculations of cluster diameter evolution with redshift¹⁶ indicate that distant rich clusters should be smaller than at the current epoch, this poses no difficulty in the present case as these evolutionary effects become noticeable only at larger redshift. As the velocity dispersion in rich clusters of galaxies is as large as $1,500 \text{ km s}^{-1}$, one might worry that the observed small limit on the velocity difference between 3C345 and 1641+3998 ($\leq 300 \text{ km s}^{-1}$) implies a cluster velocity dispersion grossly inconsistent with this norm. However, this proves not to be the case. If for simplicity, the radial velocities of the cluster galaxies are assumed to have a gaussian distribution, then we can convert the observed difference between two members into 90% confidence limits on the actual dispersion. A 300 km s^{-1} difference between two cluster members implies a cluster dispersion in the range $100\text{--}3,000 \text{ km s}^{-1}$ (90% confidence). Thus, perhaps surprisingly, the small observed velocity difference is not an awkward constraint on this model. The virial mass implied for the cluster by this allowed range of dispersions is in the range $10^{13.5}\text{--}10^{16.5} M_{\odot}$, again, not restrictive, but consistent with the parameters of known rich clusters.

Our failure to detect X-ray emission from this putative cluster is not necessarily surprising, but means that the cluster has an X-ray luminosity less than about half that of the Coma cluster (the X-ray bright core of which would appear point-like at this distance). However, Einstein observations (F. Abramopoulos, personal communication) indicate this is not unusual even for rich clusters. In addition, if 3C345 lies close to the centre of the cluster, then the possibility of source confusion implies that the cluster will be even more difficult to detect in X rays.

Our hypothesis can be tested in that the normal galaxies in this putative cluster are expected to appear at $V=22\text{--}23$ and thus be directly observable. Good quality 4-m plates of the field obtained by Kinman do show numerous faint galaxies nearby. However, the space density of clusters at this faint apparent magnitude is large enough that two or three such clusters are expected on every high latitude 4-m plate¹⁷. Thus the prob-

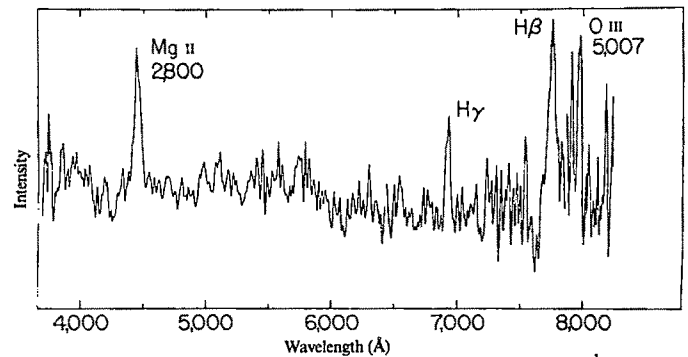


Fig. 2 The spectrum of the quasar 1641+3998, obtained with the Lick 3-m Shane reflector, with $20\text{-}\text{\AA}$ spectral resolution. The data have been converted to relative fluxes above the atmosphere through observation of spectrophotometric standard stars, and the most prominent emission features are identified.

ability of a chance superposition of a faint, unrelated cluster on this pair of QSOs is not negligible, and radial velocities of the galaxies will be necessary to verify the proposed association.

Attention has recently been drawn to other close groupings of QSOs with similar redshifts¹⁸⁻²², and diverse interpretations involving clusters, superclusters and non-cosmological redshifts have been advanced. Whether these pairs, including the one discussed here, represent clusters or superclusters is of secondary importance to the broader conclusion that QSOs exhibit similar clustering characteristics as galaxies. The pair discussed here, being at low redshift, should be suitable for future research.

We thank Drs W. Reich, F. Abramopoulos and L. Lucy for useful discussions and correspondence, and Dr T. Kinman for permission to reproduce his 4-m plate. This work has been supported by the NSF and NASA; B. M. is an Alfred P. Sloan Foundation Research Fellow.

Received 12 January; accepted 27 February 1981.

1. Gunn, J. E. *Astrophys. J. Lett.* **164**, L113-118 (1971).
2. Stockton, A. *Astrophys. J.* **223**, 747-757 (1978); *Nature*, **274**, 342-343 (1978).
3. Wyckoff, S., Wehinger, P. A., Spinrad, H. & Boksenberg, A. *Astrophys. J.* **240**, 25-31 (1980).
4. Stockton, A. in *IAU Symp.* No. 92, 89-97 (1980).
5. Giacconi, R. et al. *Astrophys. J.* **230**, 540-550 (1979).
6. Chanan, G. A., Margon, B. & Downes, R. A. *Astrophys. J. Lett.* **243** (in the press).
7. Ku, W. H.-M., Helfand, D. J. & Lucy, L. *Nature* **288**, 323-328 (1980).
8. Robinson, L. B. & Wampler, E. J. *Publ. astr. Soc. Pacif.* **84**, 161-166 (1972).
9. King, I. R. & Raff, M. I. *Publ. astr. Soc. Pacif.* **89**, 120-121 (1977).
10. Burbidge, E. M. *Astrophys. J.* **142**, 1674-1676 (1965).
11. Lynds, C. R., Stockton, A. N. & Livingston, W. C. *Astrophys. J.* **142**, 1667-1673 (1965).
12. Walsh, D., Carswell, R. F. & Weymann, R. J. *Nature* **279**, 381-384 (1979).
13. Weymann, R. J. et al. *Nature* **285**, 641-643 (1980).
14. Stute, U., Reich, W. & Kalberla, P. M. W. *Astr. Astrophys. Suppl.* (in the press).
15. Zamaroni, G. et al. *Astrophys. J.* (in the press).
16. Hickson, P. *Astrophys. J.* **217**, 964-975 (1977).
17. Hoessel, J., Gunn, J. & Oke, J. B. (in preparation).
18. Arp, H., Sulentic, J. W. & di Tullio, G. *Astrophys. J.* **229**, 489-495 (1979).
19. Hazard, C., Arp, H. C. & Morton, D. C. *Nature* **282**, 271-272 (1979).
20. Arp, H. & Hazard, C. *Astrophys. J.* **240**, 726-736 (1980).
21. Burbidge, E. M., Junkkarinen, V. T., Koski, A. T., Smith, H. E. & Hoag, A. A. *Astrophys. J. Lett.* **242**, L55-L57 (1980).
22. Oort, J. H., Arp, H. & de Ruiter, H. Preprint (Leiden University).

Random alignment of quasars

M. G. Edmunds & G. H. George

Department of Applied Mathematics and Astronomy,
University College, PO Box 78, Cardiff CF1 1XL, UK

Arp and Hazard¹ have reported several peculiar alignments of quasars, alignments which are exact to the limit of resolution on a Schmidt plate (2 arc s). We report here our investigation of the claim that the probability of such alignments occurring purely by chance is exceedingly small. Because the correct statistical analysis is not immediately obvious, we have performed a series of computer experiments to investigate how often such alignments will occur by chance.

Several trials were made in which 300 quasars were placed at random in two cartesian coordinates on a square, and each triplet of points whose overall linear extent did not exceed $1/7$ of the square side was examined for a possible alignment. Values of P , the deviation of the intermediate quasar from the line joining the other two members of the triplet, were determined. An example of the peculiar configurations that can arise by chance is shown in Fig. 1, taken from one of the trials.

The scale on the square depends entirely on the assumed density of quasars. Observational estimates² suggest that the density of quasars down to the twentieth magnitude is between 5 and 10 quasars per square degree. We have chosen a (possibly conservative) estimate of 6 quasars per square degree, (the value assumed by Arp and Hazard). On this scale, the whole square corresponds to approximately 50 square degrees of sky.

As the positions of all the quasars on the square (once placed) are known to very high accuracy, P is also known to high accuracy (corresponding to a resolution of a fraction of an arc s). In observations on a Schmidt plate, however, each image can be resolved only to 2 arc s, resulting in an uncertainty in alignment of somewhat more than 2 arc s. We have therefore chosen 2.4 arc s as a limit for P , below which we regard an alignment to be good. The number of alignments found (classified into various ranges of the value P up to 6 arc s) are shown in Table 1, the size

Table 1 The distribution of P (the displacement of the intermediate quasar of each triplet from a perfect alignment) for outer quasar separations $\leq 1^\circ$

P (arc s)	Mean of 16 runs	Largest no. in a single run	Smallest no. in a single run
0-0.24	1.6	4	0
0.24-1.20	7.4	15	4
1.20-2.40	7.9	14	4
2.40-3.60	7.4	19	2
3.60-4.80	9.4	16	5
4.80-6.00	8.2	12	5

limit on allowed alignments corresponding to 1° . Table 2 shows the number of alignments $P < 6$ arc s as a function of the separation d of the outer quasars of the alignment. χ^2 -tests indicate that these distributions are consistent with a uniform distribution in P , and a cubic distribution in d . Certain random alignments would, of course, run off the edge of the square, and thus be lost. This edge effect also occurs on Schmidt plates, but we believe that the proportion of alignments lost by this effect does not vary by more than about 10% between the square ($7^\circ \times 7^\circ$) and a typical Schmidt plate ($6^\circ \times 6^\circ$).

An estimate of alignment probabilities can be made by the following simple argument, although experience of Bertrand's paradox³ warns that such an estimate may only be correct to an order of magnitude. Let the number density of quasars be n per unit area, and consider any one quasar. Between a distance x and $x + dx$ of this quasar will lie $2\pi nx \, dx$ other quasars. A line between the first quasar and any one of these others will be of length x , and any quasars which will form triplets with distance P or less away from this line must lie within a rectangular area $2xP$, and the number of such quasars will be $2xnP$. Thus the number of alignments with the original quasar at one end is

$$\int_0^d 4\pi x^2 n^2 P \, dx = \frac{4\pi}{3} d^3 n^2 P$$

where d is the largest outer quasar separation we consider. For a plate of area A , the total number of quasars is nA , so neglecting edge effects and with a factor of $\frac{1}{2}$ to avoid counting each alignment twice, the total number of alignments N where the third quasar lies within a distance P of the line joining the other two is

$$N = \frac{2\pi}{3} A d^3 n^3 P$$

Table 2 The distribution of d (the separation of the outer two quasars of any alignment) for which $P < 6.0$ arc s

d (arc min)	Mean of 16 runs	Largest no. in a single run	Smallest no. in a single run
0-10	0.4	2	0
10-20	1.3	3	0
20-30	3.9	7	2
30-40	7.6	13	3
40-50	11.5	17	5
50-60	17.3	27	12

For comparison with the computer simulation we note $N \propto P$ and $N \propto d^3$ (as found); and if $A = 50$ square degrees, $n = 6$ per square degree, $P = 2.4$ arc s $= 6.7^\circ \times 10^{-4}$, $d = 1^\circ$ we gain $N = 15$, in excellent agreement with Table 1.

We have not considered the redshifts of the quasars. To do so would require detailed knowledge of the redshift frequency distribution, which is not well determined. Also, the cosmological interpretation of the quasar phenomenon allows for the redshift distribution to be quite independent of the distribution of quasars in the sky. Hence we could superimpose the redshifts on the existing data once the redshift distribution is sufficiently well known. Of course, one might expect one-third of the alignments to have the central quasar of lower redshift than the two outer ones. Also important may be observational selection of a central quasar on the basis of a bright magnitude, which would affect comparison with alignment probability estimates. However, for the present purposes we neglect magnitude effects (which are difficult to quantify and will be correlated with the redshift distribution).

Our computed results give an indication of how many good alignments might be expected on 50 square degrees of sky, and hence we would expect 11 alignments good to 2.4 arc s and of overall size less than a degree to appear on a typical Schmidt

F ⊙

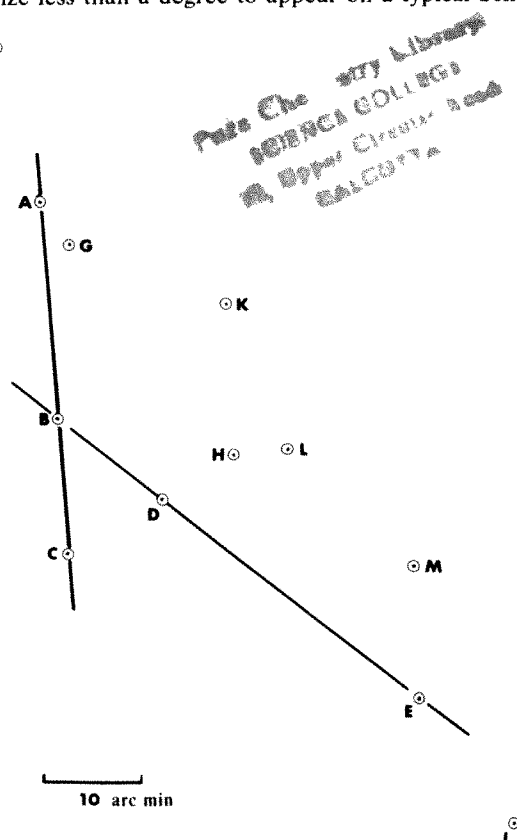


Fig. 1 An interesting square of sky (from random distribution of quasars). A, B and C are aligned to less than 0.2 arc s and B, D and E to 1.0 arc s. Also note that F, G, D and H, E, J form very good alignments (to within 2.0 arc s).

plate. It is debatable whether these are too loose criteria, and comparison with Arp and Hazard's configuration depends critically on the assumed values of n , d and P . It is particularly difficult to quantify d , the overall separation of triplets searched for. We therefore also consider a more restrictive case, and take the largest alignment found as d . Their four quoted alignments have $d < 30$ arc min, and should it be found that $P \leq 1$ arc s, then for $n = 6$ per degree we estimate an expected number of alignments $N = 0.5$ per Schmidt plate, somewhat relaxed to $N = 2.5$ if $n = 10$. Arp and Hazard's find of four alignments (albeit on an incomplete search) lies between our loose and restrictive estimates and is anyway not too great a fluctuation difference from the restrictive one. We suggest that the evidence is not strong enough to require a ballistic theory to explain their observed configuration, it could well have arisen by chance. *A posteriori* statistics on a single configuration is never satisfactory, and a more definitive test might be made by identifying all probable quasars on the plate and testing the properties of their mutual alignments against the statistics for random distribution developed above.

We thank the director and staff of University College Computing Centre for computing facilities, Professor J. V. Narlikar for discussion, and particularly Dr B. F. Schutz for originally suggesting the use of random maps.

Received 27 January; accepted 2 March 1981.

1. Arp, H. & Hazard, C. *Astrophys. J.* **240**, 726 (1980).
2. Savage, A. & Bolton, J. G. *Mon. Not. R. astr. Soc.* **188**, 599 (1979).
3. Kendal, M. G. & Moran, P. *Geometrical Probability* (Griffin, London, 1963).

Precursor active galaxies and the cosmic X-ray background

Elihu Boldt & Darryl Leiter

Laboratory for High Energy Astrophysics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

Data on the isotropic extragalactic X-ray background (>3 keV) obtained with the HEAO 1 A2 instrument¹ yield a spectrum over the energy band of maximum flux that is well described by a thermal model involving an optically thin hot plasma at a temperature of a half-billion degrees ($kT = 40 \pm 5$ keV). The spectra for known strong X-ray sources in the present epoch (active galaxies, clusters of galaxies) obtained with the same HEAO 1 instrument^{2,3} indicate that such objects at high redshifts cannot account for most of this background. On the other hand, deep-field exposures to soft X-rays (<3 keV) with the HEAO 2 Einstein Observatory⁴ suggest that discrete sources could be a sufficient explanation. This implies either (1) that these dim sources are generally unrelated to relatively bright nearby objects or (2) that they represent an earlier stage in their evolution. Short-lived thermal X-ray emission by galactic winds peculiar to an earlier epoch has been suggested⁵ as an explanation of the first kind, but the need for postulating new sources can be obviated if, as Leiter has argued⁶, prolific X-ray sources evolve in connection with active galactic nuclei involving massive compact objects. Here, we pursue the suggestion by Carr⁷ that most of the cosmic X-ray background (CXB) arises from thermal X-ray emission associated with black hole accretion disks for compact objects formed before galaxies but we favour a model where the onset of significant X-rays takes place within the epoch of galaxy formation at $z \leq 4$, when accretion disks are created by the interaction between massive black holes of pregalactic origin and young galaxies.

The spectral characteristics of the composite X-ray flux associated with the extragalactic sky is shown in Fig. 1. An extrapolation of power law (B) to the γ -ray regime ($> \frac{1}{2}$ MeV) suggests that such sources could well account for most of the

observed cosmic γ -ray background^{8,9}. But, as shown in Fig. 1, they can account for a relatively small part of the CXB.

For a survey limit of 2.6×10^{-14} erg cm⁻² s⁻¹ (1–3 keV), the deep exposures with HEAO 2 Einstein Observatory⁴ yield 19 ± 8 extragalactic sources deg⁻², comparable with the number of quasars¹⁰ brighter than $B \approx 20$ mag. Assuming that these sources have power-law X-ray spectra of energy index $\alpha = 0.7$ and that the slope a of the log N –log S relation at the survey limit and above is greater than unity, then these observed sources account for a percentage of the CXB which, at 3 keV, is given by $(a/(a-1)) (7.1 \pm 2.9)\%$. In Fig. 1, $a = 1.5$ (see ref. 4). A subsequent direct determination¹¹ of the log N –log S relation for these dim sources with HEAO 2, however, indicates that the slope is steeper and could be closer to $a = 2$. Yet, the log N –log S relation for quasars must flatten¹² as the number brighter than $B = 22.5$ mag is ≤ 50 deg⁻². The assumption that most of the CXB arises from quasars apparently gives a paradox for the required source count. Extrapolating the HEAO 2 source count through a log N –log S relation of constant slope a , for objects dimmer than the survey limit, we obtain that the number of sources required to account fully for the background is $\sim 2 \times 10^3$ sources deg⁻² for $a = 1.5$ and $\sim 10^3$ sources deg⁻² for $a = 2$. These estimates are an order of magnitude larger than the number of usual quasars considered which suggests that the onset of thermal X-rays precedes the formation of non-thermal radiating components.

The thermal-type spectrum observed for the CXB can be associated with two kinds of black-hole accretion disk radiation mechanisms: (1) bremsstrahlung from an optically thin hot plasma of the disk^{7,13}; (2) unsaturated Comptonization¹⁴ of soft photons by hot thermal electrons in the inner region of the disk, optically thin to absorption, where e^+e^- pair production cooling effects¹⁵ can limit the electron temperature (T_e) to $T_e \approx 2 \times 10^9$ K for a source luminosity close to the Eddington limit. The consistency of our model requires that the main sources of the CXB radiate at a level close to the Eddington luminosity limit. As case (1) is constrained to luminosities $\leq 1\%$ of this limit¹³, we concentrate on case (2). For reasonable sizes of the emitting hot optically thin inner disk region¹⁴ ($\leq 10^2$ GM/c²), e^+e^- cooling¹⁵ provides a thermostat whereby $kT \leq 0.2$ MeV, implying that $z \leq 4$ for most such sources of the CXB. The same condition applies to case (1) as the Gaunt factor for optically thin thermal bremsstrahlung emission at $kT \geq 0.2$ MeV would fail to provide an acceptable fit to the CXB spectrum observed (R. Shafer, personal communication). This suggests that there is no accretion disk during a black-hole accretion phase earlier than that corresponding to $z \approx 4$. For photon energies much less than the observed redshifted value of $kT \sim 40$ keV the CXB exhibits an effective energy spectral index^{1,3} $\alpha \approx 0.4$. In this restricted regime, the CXB spectrum may be synthesized by a superposition of sources having power-law spectra with $\alpha \approx 0.4 \pm 0.1$ (G. DeZotti, personal communication). For a source temperature fivefold higher than that observed for the CXB (that is, $(1+z) 40$ keV = 200 keV), case (2) provides acceptable spectral indices for values of the Kompaneets parameter ($y \approx 1-2$)

Table 1 PAG sources of the CXB

Central mass $M_0 \leq 1$ ($M_0 = 10^{-9} M/M_\odot$)
Black-hole specific angular momentum $0 \leq a/M < 0.998$ ($a = J/M$)
Luminosity (10^{47} erg s ⁻¹) (0.4–1.3) M_0
Number (4–8) $\times 10^7$
Thermal characteristics
Temperature: 2×10^9 K
Size: $\leq 10^2$ GM/c ² = $(2 \times 10^{16} M_0)$ cm
Emission mechanism: unsaturated comptonization (with Kompaneets parameter $y = 1-2$)
Redshift $z = 4$; $\Delta z = 0.3$ ($q_0 = 0.01$) or $\Delta z = 0.4$ ($q_0 = 0.5$)
Composite luminosity (10^{44} erg s ⁻¹) [*] = 1.4 ($q_0 = 0.01$) or 0.3 ($q_0 = 0.5$)
Co-moving density (10^{-4} Mpc ⁻³) [*] = 1 ($q_0 = 0.01$) or 7 ($q_0 = 0.5$)

^{*} For $h = H_0/(100 \text{ kms}^{-1} \text{ Mpc}^{-1}) = 0.5$ where H_0 = Hubble constant and q_0 = deceleration parameter.

characteristic of the model¹⁴. For photon energies $E \approx kT$, all these power-law spectra are modified by the required exponential tail-off: $\exp(-E/kT)$.

We start with Carr's suggestion that black holes are generated and evolve towards masses up to $\sim 10^9 M_\odot$ by pregalactic processes beginning at $z > 100$. However, the thermostat associated with the temperature observed for the CXB would imply that these black holes develop significant accretion disks at $z \leq 4$. Before the formation of these disks, random pregalactic processes which involve spherical accretion are not expected to be efficient for X-radiation¹⁶. Therefore, the X-radiation associated with accretion must be relatively unimportant until galaxy formation processes supply the matter associated with an ample accretion disk. Furthermore, pregalactic processes associated with random accretion should not cause appreciable spin-up of the black holes because the capture cross-section for negative angular momentum particles (opposite to the hole's spin) is greater than that for positive angular momentum¹⁷. Beginning at $z \leq 4$, within the epoch of galaxy formation, these massive black holes acquire significant accretion disks for the first time and become the 'seeds' which distinguish precursor active galaxies (PAG) from galaxies which will evolve normally. These PAG sources of the CXB postulated here are characterized in Table 1. In this model, these PAG objects would be the thermal sources of the sort required for the CXB, but necessarily cooler than previously considered ($kT \leq 0.2$ MeV). In the process of disk accretion on a time scale of $\sim 10^8$ yr, however, the central massive black hole of a PAG could increase in mass by a factor of ~ 1.5 and be spun-up¹⁸ to a 'canonical Kerr hole'. By this time black hole ergospherically induced²⁰ electron-positron injection^{6,19,33} could trigger and surge electromagnetic disk-dynamo mechanisms^{21,22}. Furthermore, the efficiency for energy transfer to the relativistic particles required for non-thermal radiation processes (such as synchrotron) could by then begin to increase appreciably, as the size of the acceleration region grows^{23,24}. Hence, during transition, there may be some young active galaxies which are not yet sufficiently luminous non-

thermal disk-dynamo synchrotron sources to be readily identified as such while their thermal disk X-ray emission is still large.

The spin-up of a black hole due to a radiating accretion disk reaches an equilibrium 'canonical' value¹⁸ ($a/M \approx 0.998$) when the hole's mass increases by $\sim 1.5 M$ (initial). Assuming that the radiation efficiency (ϵ) is constant and that the luminosity (L) is proportional to the Eddington limit (L_{Edd}), this increase of mass is achieved in a time (Δt) after the formation of the accretion disk, given by

$$(\Delta t) = 4 \times 10^8 \epsilon (1 - \epsilon)^{-1} R^{-1} \text{ yr} \quad (1)$$

where $R \equiv L/L_{\text{Edd}}$. In such conditions the black hole mass and size increase exponentially with a characteristic growth time given by equation (1). More realistically, ϵ changes during the accretion process. Initially, when $\epsilon \approx 0.06$, this growth time could be as short as 3×10^7 yr (for $R \approx 1$). At spin-up¹⁸, however, $\epsilon \approx 0.3$ and the characteristic time for further growth would be at least an order of magnitude longer (with $R < 1$). If the accretion rate is then 'starved' by a high angular momentum configuration of remaining gas, such as in a spiral galaxy, we might expect $R \ll 1$ and that any further growth within a Hubble time ($\sim 10^{10}$ yr) is relatively minor. The amorphous nature of the spiral galaxies associated with Seyfert nuclei may well be caused by a process which slowly feeds material from outer regions into the centre²⁵.

By the time of spin-up, given by equation (1), the innermost stable orbit of the accretion disk has amply penetrated the ergosphere²⁶, and electron-positron injection processes⁶ associated with a canonical Kerr black hole are 'switched on'. The time (Δt) to reach this state is the effective lifetime of the PAG phase. This canonical state of spin is insensitive to small ($< 10\%$) fluctuations in the matter accretion rate and the counter-torques due to radiation¹⁸. In this context, the evolution of a PAG into a system involving a massive canonical Kerr hole may be viewed as an essentially irreversible process. Such behaviour is of the type needed for describing the observed evolution²⁷ whereby the sources of the CXB (for example, PAG objects at $z \leq 4$) lead directly to the active galaxies observed at later epochs. As a broad band of non-thermal radiation can eventually be generated at a relatively high efficiency²⁸ (at the expense of weakened thermal disk emission^{21,22}), the thermal disk component could be masked for many of the brightest active galaxies, but might sometimes be observable. In this connection, the spectrum measured for the quasar 3C273 is consistent²⁹ with a substantial contribution from a thermal X-ray component at $T \geq 10^9$ K, while the broad-band X-ray spectrum of QSO 0241+622 is incompatible with such a feature³⁰. The PAG objects associated with Seyfert nuclei might involve galaxies where the amount of gas available for accretion from a low angular momentum configuration (for example, in a central bulge) is not large compared with that of the initial black hole. In this situation (such as one involving a spiral galaxy) we expect that a luminosity as high as 10^{47} ergs s^{-1} (Eddington limit for $10^9 M_\odot$) could not last much beyond spin-up. For appreciably longer time scales we assume $R \ll 1$; the resulting active galaxies could thereby survive to be observed in the present epoch (for example, as Seyfert galaxies) with much lower luminosities but with non-thermal spectra extending well into the γ -ray band.

The thermal component of the CXB corresponds to a surface brightness of 2.0×10^{-7} erg $(\text{cm}^2 \text{ s sr})^{-1}$. For a Friedman cosmology with $q_0 \leq 0.5$, $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ and with most sources of the CXB at $z \approx 4$, the bolometric luminosity for the sum of all these sources, in their proper frame, is $\Sigma L \geq 8 \times 10^{52} h^{-2} \text{ erg s}^{-1}$. The number of sources required for the CXB is $N(\text{CXB}) \approx (4-8) \times 10^7$, so the implied average bolometric luminosity per source ($\bar{L} \geq 10^{45} h^{-2} \text{ erg s}^{-1}$) is readily associated with the PAG objects postulated here as the limit $R > 0.3$ is required to constrain the PAG lifetime from being much greater than $\sim 10^8$ yr (see equation (1)). If the central black hole were initially much less massive than $\sim 10^7 M_\odot$, then it would take substantially more than 10^8 yr (beyond spin-up) to

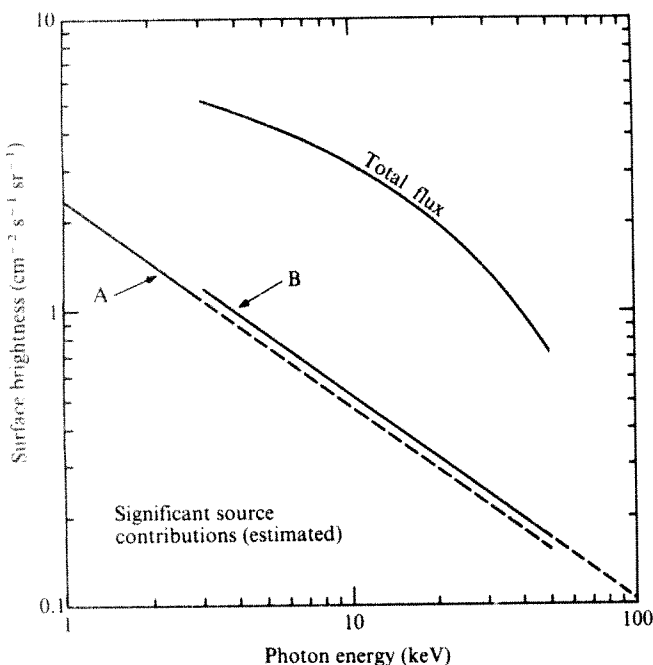


Fig. 1 Spectral density for the surface brightness of the extragalactic X-ray sky as a function of photon energy. The 'total flux' curve is the best-fit thermal spectrum ($kT = 40$ keV) for the background measured by Marshall *et al.*¹. The power-law A represents the composite flux due to sources detected above the survey limit in the HEAO 2 deep exposures by Giacconi *et al.*⁴. The power-law 'B' represents the composite flux from active galaxies with $z \leq 1$, based on the local luminosity function determined³ with HEAO 1 A2 and assuming no evolution, with $q_0 = 0$. The dashed lines are extrapolations.

grow towards a mass where the luminosity could reach the level required here for sources at $z \approx 4$. Hence, such objects would not qualify as PAG sources of the CXB.

The total number of PAG objects $N(\text{PAG})$ making up the CXB is fixed by the model in terms of the local density of active galaxies. A fundamental aspect of this model is that the density of PAG objects in co-moving coordinates remains constant during the time (Δt) of spin-up and that the co-moving density of active galaxies after production equals that of the associated PAG objects. Assuming that the PAG population is produced at a look-back time corresponding to $z = 4$, $N(\text{PAG})$ is given by

$$N(\text{PAG}) = (1.4 - 8.9) \times 10^8 \phi \Delta \tau h^{-3} \quad (2)$$

where the limits indicated correspond to $q_0 = 0.5$ and $q_0 = 0.01$ respectively, $\Delta \tau$ (units of age of the universe) is the look-back temporal interval of the PAG phase, and ϕ is the co-moving density of PAG objects (in units of 10^{-4} Mpc^{-3} , the value measured²⁵ for the local density of Seyfert galaxies and consistent with the local luminosity function for X-ray active galaxies³ in the luminosity range $\sim 10^{42} - 10^{45} \text{ erg s}^{-1}$ (3–50 keV)). For $\phi \geq 1$ and $\Delta \tau \approx 10^{-2}$ (corresponding to $\sim 10^8 \text{ yr}$) equation (2) indicates that the number of PAG objects could well match the number of sources required for the CXB: $N(\text{CXB}) \approx (4 - 8) \times 10^7$. The total mass involved in all the pregalactic black holes that precipitated these sources would be less than 10^{-4} of the known mass in galaxies. When most PAG objects are within $\Delta \tau = 10^{-2}$ at $z \approx 4$ then $\Delta z < 8 \times 10^{-2}(1+z)$, ensuring that the fractional spread in observed temperatures due to redshift would be $< 8\%$, compatible with the CXB spectral data.

We conclude that most of the CXB could be generated by distant sources that are not usual active galaxies or quasars, and that they must have a relatively low ratio of optical to X-ray luminosity (that is dominantly X-ray objects). PAG objects at $z \leq 4$ involving a massive ($\leq 10^9 M_\odot$) black hole surrounded by a hot optically thin accretion disk radiating near the Eddington luminosity limit with an effective temperature $\sim 2 \times 10^9 \text{ K}$ would meet this condition and yield the correct CXB spectrum and intensity. Furthermore, this model has the advantage of providing an evolutionary track which links both the CXB and the non-thermal γ -ray background directly to active galaxies, implying that massive ($\leq 10^9 M_\odot$) black holes are indeed associated with their internal dynamics (as suggested by X-ray observations³¹ involving temporal variability). While the PAG sources of the CXB postulated here would have low optical continuum emission relative to quasars, strong emission lines comparable with those of quasars generated from a broad-line region photoionized³² by the central X-ray source could lead to a direct identification of these objects. As the average X-ray flux expected from a PAG is an order of magnitude below the survey limit for the HEAO 2 Einstein Observatory, they would not be well represented in the 'complete' sub-sample observed, but some of the dimmest unidentified objects detected should be considered candidates.

We thank Richard Mushotzky for valuable discussions. D.L. is a NAS-NRC research associate.

Received 8 September 1980; accepted 7 January 1981.

1. Marshall, F. *et al. Astrophys. J.* **235**, 4 (1980).
2. Boldt, E. *Bull. Am. Ass. Sci.* **11**, 651 (1979).
3. Boldt, E. *Comments Astrophys.* (in the press).
4. Giacconi, R. *et al. Astrophys. J. Lett.* **234**, L1 (1979).
5. Bookbinder, J., Cowie, L., Krolik, J., Ostriker, J. & Rees, M. *Astrophys. J.* **237**, 647 (1980).
6. Leiter, D. *Astr. Astrophys.* **89**, 370 (1980).
7. Carr, B. *J. Nature* **284**, 326 (1980).
8. Bignami, G., Fichtel, C., Hartman, R. & Thompson, D. *Astrophys. J.* **232**, 649 (1979).
9. Fichtel, C. & Trombka, J. *Gamma-Ray Astrophysics, New Insights into the Universe* Ch. 9 (NASA-RP-007 in the press).
10. Wills, D. *Phys. Scr.* **17**, 333 (1978).
11. Murray, S. *Bull. Am. Ass. Sci.* **11**, 642 (1979).
12. Bahcall, J. N. & Soneira, R. M. *Astrophys. J. Lett.* **238**, L17 (1979).
13. Payne, D. & Eardley, D. *Astrophys. Lett.* **19**, 39 (1977).
14. Shapiro, S., Lightman, A. & Eardley, D. *Astrophys. J.* **204**, 187 (1976).
15. Liang, E. P. T. *Astrophys. J.* **234**, 1105 (1980).
16. Brinkmann, W. *Astr. Astrophys.* **85**, 146 (1980).
17. Ruffini, R. & Wheeler, J. in *Neutron Stars, Black Holes and Binary X-Ray Sources* (eds Gursky, H. & Ruffini, R.) 389 (Reidel, Dordrecht, 1975).
18. Thorne, K. *Astrophys. J.* **191**, 507 (1974).
19. Leiter, D. & Kafatos, M. in *Proc. La Jolla Institute Workshop on Particle Acceleration Mechanisms in Astrophysics*, 419 (American Institute of Physics, New York, 1979).

20. Piran, T., Shaham, J. & Katz, J. *Astrophys. J. Lett.* **196**, L107 (1975).
21. Blandford, R. D. in *Active Galactic Nuclei* (eds Hazard, C. & Mitton, S.) 241 (Cambridge University Press, 1979).
22. Lovelace, R., MacAuslan, J. & Burns, M. in *Proc. La Jolla Institute Workshop on Particle Acceleration Mechanisms in Astrophysics* 399 (American Institute of Physics, New York, 1979).
23. Vestrand, W., Scott, J., Marsher, A. & Christiansen, W. *Astrophys. J.* **245** (in the press).
24. Cavaliere, A. & Morrison, P. *Astrophys. J. Lett.* **238**, L63 (1980).
25. Simkin, S., Su, H. & Schwarz, M. *Astrophys. J.* **237**, 404 (1980).
26. Bardeen, J. in *Black Holes* (eds DeWitt, C. & B. S.) 225 (Gordon and Breach, New York, 1973).
27. Boldt, E. *Bull. Am. Ass. Sci.* **12**, 495 (1980).
28. Leiter, D. *Bull. Am. Ass. Sci.* **12**, 496 (1980).
29. Jones, T. *Astrophys. J.* **233**, 796 (1979).
30. Worrall, D. *et al. Astrophys. J.* **232**, 683 (1979).
31. Worrall, D., Boldt, E., Holt, S. & Serlemitsos, P. *Astrophys. J.* **240**, 421 (1980).
32. Marshall, N., Warwick, R. S. & Pounds, K. A. *Mon. Not. R. astr. Soc.* (submitted).
33. Grindlay, J., Steiner, J., Forman, W., Canizares, C. & McClintock, J. *Astrophys. J. Lett.* **239**, L43 (1980).
34. Kafatos, M. & Leiter, D. *Astrophys. J.* **229**, 46 (1979).

Measurement of ionospheric winter anomaly in Southern Hemisphere

Brian G. Ferguson

Ionospheric Prediction Service, Department of Science and the Environment, PO Box 702, Darlinghurst, 2010 Sydney, Australia

The unexpectedly high ionospheric radio-wave absorption that occurs during winter at temperate latitudes is referred to as the 'winter anomaly'. These variations have been attributed to increases in the concentration of the minor atmospheric species NO and O₂(¹Δ_g), themselves important in the formation of the D-region. This anomaly involves an order of magnitude increase of the electron number density which, in turn, leads to an increase in the absorption experienced by radio waves propagating in the D-region. The phenomenon occurs between 75 and 95 km, with the maximum enhancement near 80–83 km (ref. 1). We describe here a period of anomalous absorption observed at Camden, Australia (34°S, 151°E), during 15–17 June 1980. This anomalous event is significant because previous observations of the winter anomaly have, in general, been restricted to the Northern Hemisphere, especially Europe and because our observation was made at a lower latitude than those suggested^{2,3} as the low-latitude boundary of the winter absorption anomaly. Previous measurements in the Northern Hemisphere have placed the low-latitude boundary at about 37°N (for the 1968–69 winter²) and 35°N (for the 1971–72 winter³). The second measurement suggested an increase of 1 dB per degree of latitude above the boundary. The magnitude of the anomalous absorption in our measurements, at 34°S, was 10 dB. The magnitude of this absorption suggests either that this event was very large or that the ionospheric response is different in the two hemispheres.

The A1 (or pulse amplitude) method for measuring ionospheric absorption involves the measurement of the amplitude of radio-wave pulses reflected vertically from the ionosphere. An automatic A1 system, which uses a micro-processor for both the various system control functions and the signal processing has been described elsewhere⁴. The equipment was installed at Camden, New South Wales, in December 1979 and operates on a frequency of 1.91 MHz. A sample of the amplitude data (3-min medians) for 20 June 1980 is plotted in Fig. 1 as a function of local time (GMT + 10 h) between 08.00 and 16.00 h.

A convenient technique for characterizing the diurnal variation in the amplitude data is to fit a second-degree polynomial (quadratic) to the observations by using least-square methods. For instance, the quadratic which best fits the observations for 20 June between 08.00 and 16.00 h is shown in Fig. 1. The scatter of points about this (solid) curve is attributed mainly to either non-dissipative amplitude fluctuations (fading) or to variations in the reflection mode. In addition, solar X-ray events can cause the data points to depart significantly (by $> 10 \text{ dB}$) from the quadratic curve describing the normal daily variation in the amplitude data. The ionospheric response to X rays of short

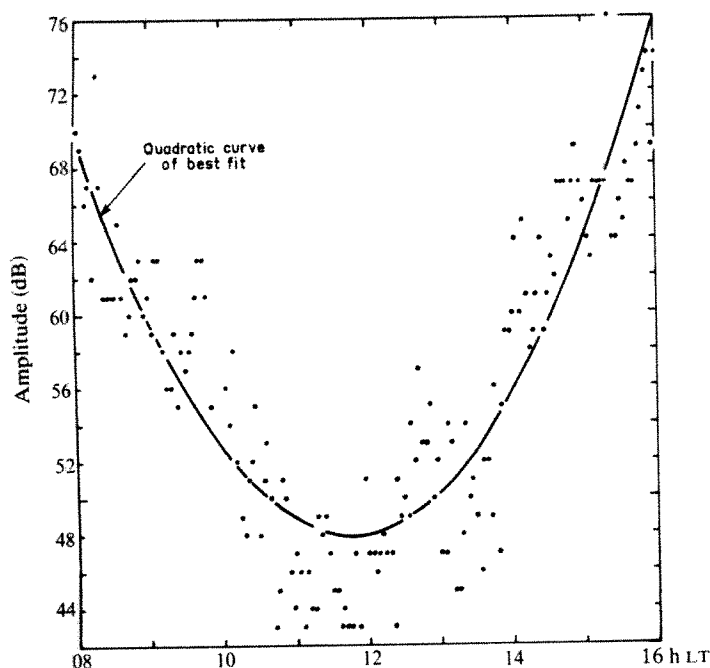


Fig. 1 A1 amplitude data (3-min medians) for the period 08.00–16.00 h, 20 June 1980. The full curve represents the second-order polynomial of best fit through the data points.

wavelengths ($<10 \text{ \AA}$) is to increase the concentration of electrons in the D-region through ionization of the neutral atmosphere. This is accompanied by an increase in the absorption suffered by radio waves as they travel through the region of enhanced electron density.

The quadratic curves of best fit (one for each day) for the periods 8–20 June (solid curves) and 4–11 July (broken curves) are superimposed in Fig. 2. A1 data for the intervening period were not available. Solar X-ray events during these periods were either of small intensity (C class events) or of small duration ($<1 \text{ h}$), which indicates that the ionospheric response was not large enough to modify significantly the coefficients of the quadratic curves of best fit. However, this was not true for the data collected on 6 July, during which a large (M9) X-ray event produced an absorption increase of at least 60 dB between 08.40 and 09.05 LT. Consequently, the A1 amplitude data associated with this event were omitted in determining the quadratic curve for 6 July.

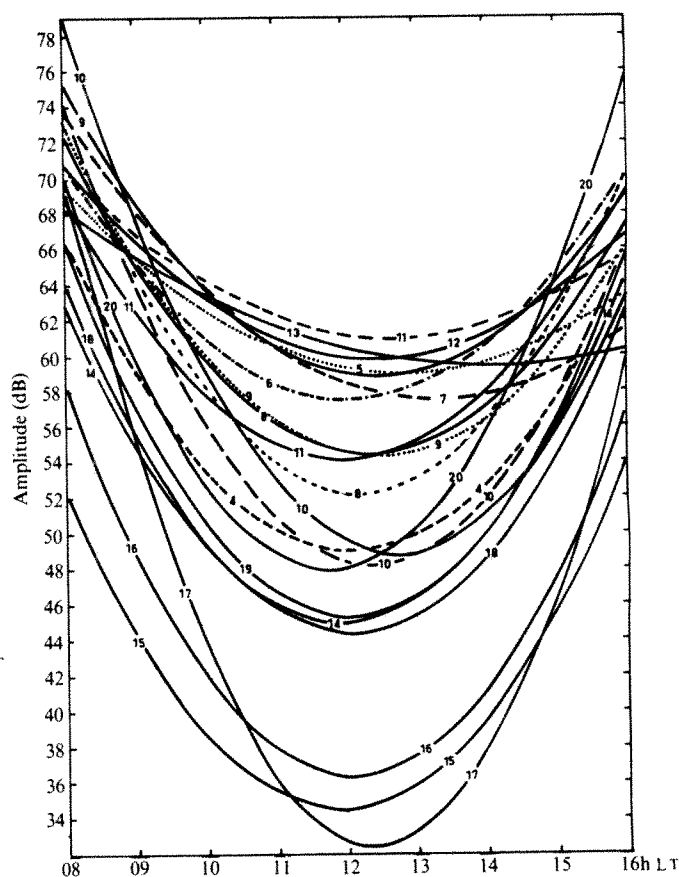


Fig. 2 Quadratic curves of best fit for the A1 amplitude data for the periods 8–20 June (solid curves) and 4–11 July (broken curves). The days that exhibit marked anomalous absorption are 15, 16 and 17 June 1980. Numbers on the curve indicate days of the month.

Figure 2 shows that the days of 15–17 June were characterized by amplitude values which were consistently below those for the other days. Thus 15–17 June are classified as days which exhibited anomalous absorption. The minimum amplitude values of the quadratic curves associated with these anomalous days are $\sim 10 \text{ dB}$ less than those for the adjacent days of 14, 18 and 19 June. However, by considering days remote from the anomalous period (say 10 June and 10 July), the anomalous absorption seems to be $\sim 15 \text{ dB}$. If the latter estimate of the magnitude of

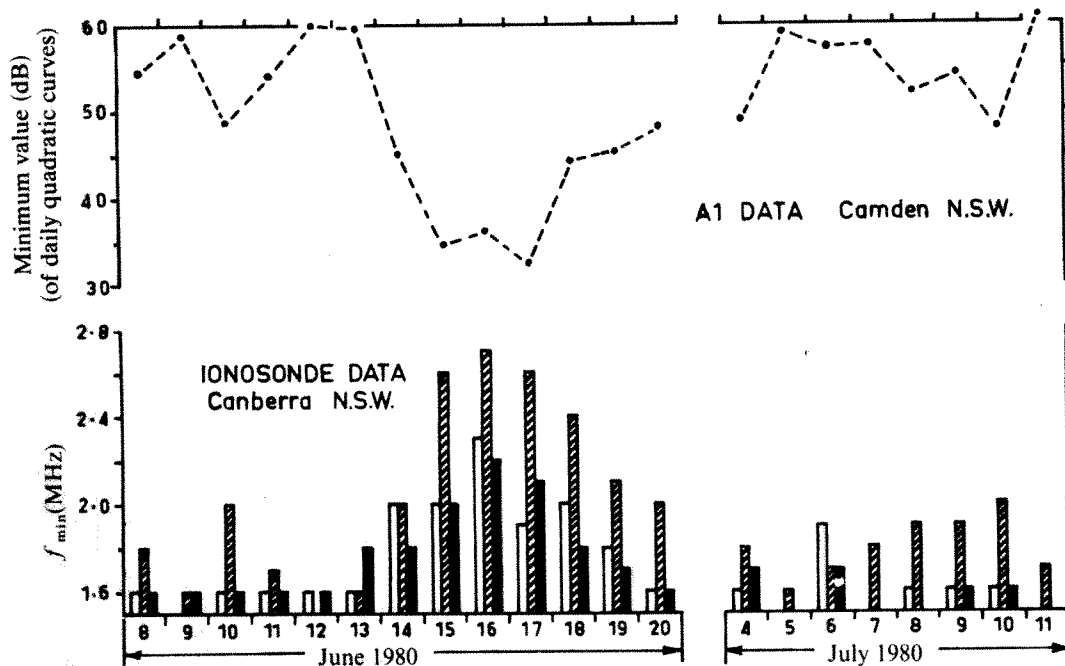


Fig. 3 Minimum values of the daily quadratic curves of best fit for the Camden A1 data, together with a histogram of the Canberra $f_{\min}$ data (median of the 3-h values), plotted for the periods 8–20 June and 4–11 July 1980. $\square$, 08.00–10.00 h; $\square$, 11.00–13.00 h; $\blacksquare$, 14.00–16.00 h LT.

the anomalous absorption is adopted, then the period associated with the winter anomaly can be extended to include 14, 18 and 19 June. Thus the anomalous absorption for 14, 18 and 19 June was ~3 dB, while for 15–17 June it was ~15 dB.

The winter anomaly in absorption also manifests itself in the Canberra (35°S, 149°E) ionosonde data. The appropriate parameter is $f_{\min}$ which is the lowest frequency at which an echo returned from the ionosphere can be detected on an ionogram. This parameter is observed to increase as absorption increases, for example, during solar X-ray flares. However, $f_{\min}$ is also affected by factors extraneous to ionospheric absorption, for instance, equipment characteristics and radio-frequency interference. Figure 3 shows the medians of the scaled $f_{\min}$ hourly values for the periods 08.00–10.00, 11.00–13.00 and 14.00–16.00 h using a histogram format for the periods 8–20 June and 4–11 July. For comparison, a graph of the minimum values of the quadratic curves of best fit for the A1 data is also shown for the same periods. The days characterized by enhanced $f_{\min}$ values coincide with days for which the absorption was anomalously high.

The above preliminary analysis indicates that a period of anomalous absorption occurred during the winter of 1980 in the Southern Hemisphere. The A1 amplitude data from Camden indicated that the anomalous absorption had a lower limit of 10 dB and may have been as high as 15 dB, depending on the reference level. Enhanced $f_{\min}$ values were recorded on the anomalous days by an ionosonde located in Canberra. The results show that an asymmetry exists between the Northern and Southern Hemispheres with respect to winter anomalous absorption. Predictions based on data from only the Northern Hemisphere were found to be inappropriate as they failed to explain the existence of anomalous absorption at Camden. This suggests that the experimental programmes aimed at collecting data on the aeronomic (physical, chemical and dynamical) behaviour of the stratosphere/mesosphere during the forthcoming Middle Atmosphere Programme (1982–85), should be conducted in both hemispheres. The results should be analysed separately (according to hemisphere) rather than combining them into one data base which would average-out any asymmetry. Also, global aeronomic models should reflect the differences that exist between the hemispheres.

I thank George Daniel for assistance with the diagrams.

Received 20 October 1980; accepted 14 January 1981.

1. Offerman, D. J. *atmos. terr. Phys.* **41**, 735–752 (1979).
2. Lauter, E. A. & Schänning, B. J. *atmos. terr. Phys.* **32**, 1619–1624 (1970).
3. Beynon, W. J. G. & Williams, E. R. J. *atmos. terr. Phys.* **38**, 423–429 (1976).
4. Ferguson, B. G. *Proc. Instn Radio Engrs Aust.* **41**, 105–108 (1980).

Inorganic lead complexation in natural seawater determined by UV spectroscopy

Robert H. Byrne

Department of Marine Science, University of South Florida, St Petersburg, Florida 33701, USA

Recent effort to construct trace metal complexation models in natural media^{1–8} is a mark of the importance of chemical form in the biogeochemical behaviour of trace metals. Unfortunately such chemical modelling is difficult in many ways. Chemical speciation models of trace metals in complex natural media such as seawater are usually constructed using stability constants determined in simple media. We describe an alternative procedure which can reduce the errors inherent in this approach. Using UV spectroscopy, we have examined lead speciation in seawater. The method is a means of describing metal speciation in seawater quantitatively using stability constants determined exclusively in seawater. Preliminary experiments indicate that the same approach can be used for other metals including copper.

Speciation models for lead in seawater (Table 1) alternatively predict: (1) that only lead carbonate complexes¹ are important; (2) that lead carbonate species are unimportant and that lead chloride complexes are dominant^{2,3}; (3) that lead hydroxide complexes are important^{4,5}; and (4) that lead hydroxide complexes are unimportant⁶. Sipos *et al.*⁵ have described two models of lead complexation in seawater. In their second model, they adopt the chloride speciation scheme of Johnson and Pytkowicz⁹ and the concentration ratio of their dominant lead species, $\text{PbOH}^+/\text{PbCO}_3^0$, diminished by a factor of 3.5 compared with the first model. Three factors explain the observed variability of models. (1) Stability constants are selected from a wide range of values. (2) Stability constants are generally determined in simple media and adapting these constants to seawater conditions introduces error. (3) Minor element speciation schemes depend on the speciation schemes adopted for the major constituents of seawater. Our approach eliminates the latter two sources of error, and provides an essential test of the various models constructed for lead speciation in seawater.

Lead chloride complexation experiments used unfiltered surface seawater (34.7‰ salinity) from the Gulf of Mexico which was enriched in lead (~7 μM in Pb(II)) with an acidic lead stock solution. The UV absorbance characteristics of these solutions were measured between 300 and 220 nm against a lead-free seawater reference solution. The test solution, housed in a 10-cm pathlength spectrophotometric cell, was acidified with varying amounts of 1 M HCl, producing a series of absorbance versus pH curves at 25 °C (Fig. 1). Cells were sealed except during measurements of pH. For each absorbance versus wavelength curve (Fig. 1) the solution pH ($-\log A_{\text{H}}$) was determined with a micro combination pH probe (Micro-electrodes Inc.) calibrated on the NBS scale. Experiments were conducted at the naturally occurring total CO_2 concentration, 2.05×10^{-3} mol total CO_2 per kg seawater. Total CO_2 and M_{CO_3} , the total carbonate ion concentration, were calculated from pH and the initial carbonate alkalinity^{10,11} using the apparent constants of Mehrbach *et al.*¹².

The isosbestic point observed at 237 nm in Fig. 1 indicates that adsorption is not significant. The strong pH dependence demonstrates that lead is not exclusively complexed with chloride ions in seawater. PbCO_3^0 and PbOH^+ are thought to be important at high pH. Varying the pH of a lead-enriched seawater solution between 3.9 and 8.2 causes molar-absorptivity changes between 220 and 300 nm which are highly dependent on the seawater's total CO_2 content. The molar-absorptivity change is $<250 \text{ M}^{-1} \text{ cm}^{-1}$ at zero total CO_2 and as large as $1,400 \text{ M}^{-1} \text{ cm}^{-1}$ at the natural CO_2 concentration, 2.05×10^{-3} mol per kg seawater. The much larger change when carbonate is present suggests that PbOH^+ is much less significant than PbCO_3^0 . Hence plots of absorbance versus pH at constant wavelength will be distinctly different for different total CO_2 concentrations, and absorbance versus $\log M_{\text{CO}_3}$ at different pH will be nearly identical (Fig. 2).

The results of five experiments were analysed quantitatively: three at total CO_2 2.05×10^{-3} mol per kg seawater and two at total CO_2 9.84×10^{-3} mol per kg seawater. Assuming PbOH^+ is insignificant compared with PbCO_3^0 , the absorbance of lead in seawater can be described by^{13,14}:

$$_{\lambda}A = \frac{_{\lambda}\epsilon_0 + _{\lambda}\epsilon_1\beta_1 \cdot M_{\text{CO}_3}}{1 + \beta_1 \cdot M_{\text{CO}_3}} \quad (1)$$

where $_{\lambda}A$ is absorbance at wavelength λ , M_{CO_3} is the sum concentration of carbonate species (such as, CO_3^{2-} , NaCO_3^- , MgCO_3^0 , and CaCO_3^0), $_{\lambda}\epsilon_1$ is absorbance due to PbCO_3^0 , $_{\lambda}\epsilon_0$ is absorbance due to the lead forms which dominate at low pH (principally Pb^{2+} , PbCl^+ , PbCl_2^0 and PbCl_3^-) and β_1 is the formation constant of PbCO_3^0 .

Lead-absorbance data were analysed at four wavelengths (240, 250, 260 and 270 nm) which did not require corrections for baseline variability with pH. Absorbance data at these

Table 1 Species composition of inorganic lead in seawater

Model-species	Ref. 2*	Ref. 3	Ref. 7	Ref. 4	Ref. 5†		Ref. 8	Ref. 6	Ref. 1
PbCO ₃ ⁰ , Pb(CO ₃ ²⁻) ₂ ²⁻ , PbCO ₃ Cl ⁻	—	<1	18	40	22	47	65	77	100
PbOH ⁺ , Pb(OH) ₂ ⁰ , PbOHCl ⁰	2	15	5	33	49	30	6	1	—
Pb ²⁺ , PbCl ⁺ , PbCl ₂ ⁰ , PbCl ₃ ⁻ , PbSO ₄ ⁰	98	84	73	27	28	23	28	21	—

Composition is given as a percentage of the total inorganic lead at pH 8. Totals <100% indicate that additional complexes contribute to the speciation scheme.

* The Morel and Morgan² model was constructed for a hypothetical system which is quite similar to seawater in ionic strength and the free concentrations of Cl^- , CO_3^{2-} and H^+ .

† The Sipos *et al.*⁵ model is constructed for pH 8.1.

wavelengths in five experiments were fit using equation (1) in a non-linear least-squares analysis^{13,14}. At 25.0 °C and 34.7% salinity:

$$\beta_1 = \frac{M_{\text{PbCO}_3}}{M_{\text{Pb}} \cdot M_{\text{CO}_3}} = (10.0 \pm 0.3) \times 10^3 \quad (\text{mol per kg seawater})^{-1} \quad (2)$$

where M_{PbCO_3} is the concentration of PbCO_3^0 and mixed ligand carbonate species^{8,15,16}, and M_{Pb} is the sum concentration of Pb^{2+} , PbCl^+ , PbCl_2^0 , PbCl_3^- and less significant species. This PbCO_3^0 formation constant enables $M_{\text{PbCO}_3}/\text{total Pb}$ at 25.0 °C

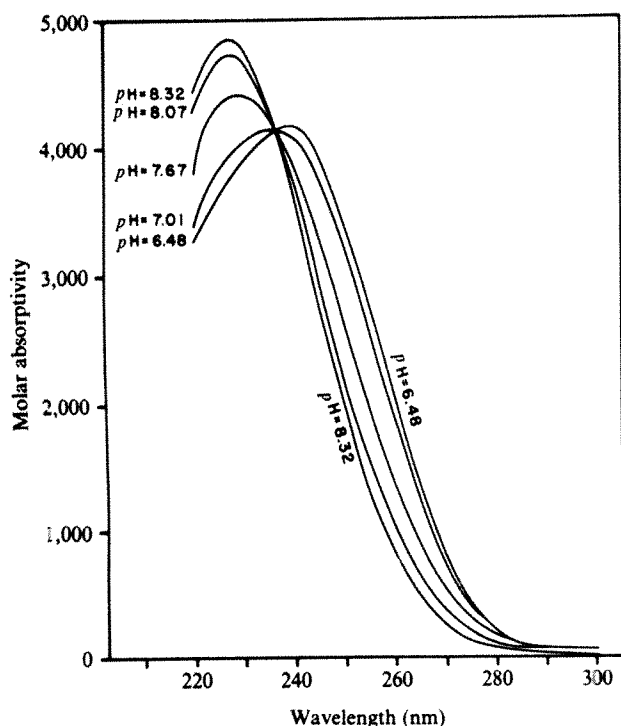


Fig. 1 Absorbance per molar cm of $\text{Pb}(\text{II})$ in natural seawater as a function of wavelength and pH. (Temperature 25.0 °C; salinity 34.7%; total CO_2 : 2.05×10^{-3} mol per kg seawater.) Absorbance data are corrected for baseline variation with pH at wavelengths <235 nm.

and 34.7% salinity to be calculated when total CO_2 and pH are known. M_{CO_3} is calculated from total CO_2 and pH (see refs 11, 12). The fraction of total lead present as PbCO_3^0 can then be determined without knowledge of the ion pairing interactions of CO_3^{2-} with the major cationic seawater constituents, and there is no formation constant variation between seawater and the test medium.

Using equation (2) at 25.0 °C, 34.7% salinity and total CO_2 of 2.05×10^{-3} mol per kg seawater it can be shown that $M_{\text{PbCO}_3}/\text{Total Pb} \geq 0.5$ for $\text{pH} \geq 7.85$. As seawater generally

contains a greater total CO_2 concentration, PbCO_3^0 seems to be the dominant form of inorganic lead in seawater.

Expansion of equation (1) to account for the possible existence of PbOH^+ , PbHCO_3^+ and $\text{Pb}(\text{CO}_3)_2^{2-}$ in the test medium produced β_1 values similar to that in equation (2). Although formation constants for the species PbOH^+ , PbHCO_3^+ and $\text{Pb}(\text{CO}_3)_2^{2-}$ were ill-defined suggesting that these species were present only in low concentration, the analyses did produce upper 95% confidence limits which are useful for estimating the importance of each species in seawater (Table 2). The formation constant determined for PbHCO_3^+ indicates that this species has little significance ($M_{\text{HCO}_3} \approx 2 \times 10^{-3}$ mol per kg seawater). From its formation constant, PbOH^+ is unlikely to constitute more than 10% of the total inorganic Pb in seawater ($\text{pH} \leq 8.3$). From the β_2 upper 95% confidence limit, $\text{Pb}(\text{CO}_3)_2^{2-}$ and PbCO_3^0 could be present at approximately equal concentrations in seawater ($M_{\text{CO}_3} \approx 3 \times 10^{-4}$ mol per kg seawater), but the existence of well defined isosbestic points makes this improbable. If $\text{Pb}(\text{CO}_3)_2^{2-}$ were present at appreciable concentrations, its molar absorbance at 237 nm would have to be identical to that of PbCO_3^0 to produce the observed isosbestic point.

The present analysis requires that a 7 μM lead enrichment would not significantly alter the inorganic complexation of lead in seawater. This is satisfied because the appropriate ligands have concentrations $\gg 7 \mu\text{M}$. The most unfavourable molar ratio, total $\text{CO}_2/\text{Pb}(\text{II})$, is >290 . Consequently, calculated values of M_{CO_3} are more affected by small experimental errors in measurement of pH than by lead enrichment, enabling this description of inorganic lead complexation to be applied to natural seawater.

Our lead carbonate formation constants are strongly influenced by the ion pairing interactions of carbonate ion and possibly chloride ion with the major cationic constituents of seawater. However, as our methods do not require any particular model of these interactions, our formation constants and conclusions are independent of these descriptive models.

Constants of the form defined by equation (2) allow the pH dependence of Pb speciation in seawater to be described without assuming that activity coefficients are constant in diverse media at constant ionic strength. Our conclusions that the dominant

Table 2 Results of formation constant determinations in natural seawater at 25.0 °C and 34.7% salinity

Species considered	Formation constants determined
PbCO_3^0	$\beta_1 = 10.0 \times 10^3$
PbCO_3^0 , PbHCO_3^+	$\beta_1 = 9.4 \times 10^3$; $\text{HCO}_3\beta = \frac{M_{\text{PbHCO}_3}}{M_{\text{Pb}}M_{\text{HCO}_3}} \leq 14.5$
PbCO_3^0 , PbOH^+	$\beta_1 = 10.3 \times 10^3$; $\beta_1 = \frac{M_{\text{PbOH}^+}A_{\text{H}}}{M_{\text{Pb}}} \leq 2 \times 10^{-9}$
PbCO_3^0 , $\text{Pb}(\text{CO}_3)_2^{2-}$	$\beta_1 = 12.3 \times 10^3$; $\beta_2 = \frac{M_{\text{Pb}(\text{CO}_3)_2}}{M_{\text{Pb}}M_{\text{CO}_3}^2} \leq 7.3 \times 10^7$

Concentrations of all chemical species except H^+ are defined on a mol per kg of seawater concentration scale. Hydrogen ion activities, A_{H} , are operationally defined on the NBS scale¹¹.

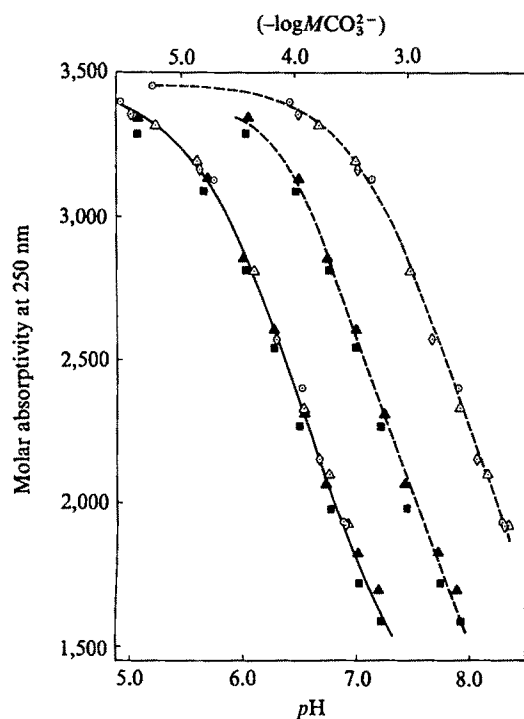


Fig. 2 Absorbance per molar cm of Pb(II) in seawater as a function of pH and $\log M_{\text{CO}_3}$ at 25.0 °C and 34.7‰ salinity. Total CO_2 is 2.05×10^{-3} mol per kg seawater ($\Delta \diamond$) or 9.84×10^{-3} mol per kg seawater ($\blacksquare$). Dashed lines are coordinated on the pH axis. The solid line is coordinated on the $\log M_{\text{CO}_3}$ axis.

inorganic species of lead in seawater are PbCO_3^0 and lead chloride species are consistent with the models of Turner and Whitfield⁸ and Zirino and Yamamoto⁶.

Other trace metals such as iron (III) and copper have intense absorbance spectra in the UV. The methods used here can be extended to investigations of the chemistry of iron and copper in natural media. Preliminary experiments indicate that studies of copper in seawater can be conducted within the natural pH range of seawater. The plasticity of the copper (II) coordination sphere¹⁷ suggests that the activity coefficients of copper species may be significantly influenced by medium composition, and hence that the 'natural medium approach' may be especially useful.

This research was supported by contract OCE-79-19397 from the NSF and by contract EE-77-S-05-5486 from the Department of Energy. I thank Bill Miller for assistance in data reduction and analysis.

Received 8 September; accepted 15 December 1980.

1. Rickard, D. T. & Nriagu, J. O. in *The Biogeochemistry of Lead in the Environment, Part A* (ed. Nriagu, J. O.) 219–285 (Elsevier, Amsterdam, 1978).
2. Morel, F. & Morgan, J. *Envir. Sci. Technol.* **6**, 58–67 (1972).
3. Dyrssen, D. & Wedborg, M. in *The Sea* (ed. Goldberg, E. D.) 181–195 (Wiley, New York, 1975).
4. Stumm, W. & Brauner, P. A. in *Chemical Oceanography* (eds Riley, J. P. & Skirrow, G.) 173–240 (Academic, New York, 1975).
5. Sipos, L., Raspor, B., Nürnberg, H. W. & Pytkowicz, R. M. *Mar. Chem.* **9**, 37–47 (1980).
6. Zirino, A. & Yamamoto, S. *Limnol. Oceanogr.* **17**, 661–671 (1972).
7. Long, A. & Angino, E. E. *Geochim. cosmochim. Acta* **41**, 1183–1191 (1977).
8. Turner, D. R. & Whitfield, M. *J. electroanal. Chem.* **103**, 61–79 (1979).
9. Johnson, K. S. & Pytkowicz, R. M. *Am. J. Sci.* **278**, 1428–1447 (1978).
10. Culbertson, C., Pytkowicz, R. M. & Hawley, J. E. *J. mar. Res.* **28**, 15–21 (1970).
11. Millero, F. J. *Geochim. cosmochim. Acta* **43**, 1651–1661 (1979).
12. Mehrbach, C., Culbertson, C. H., Hawley, J. E. & Pytkowicz, R. M. *Limnol. Oceanogr.* **18**, 897–907 (1973).
13. Byrne, R. H. & Kester, D. R. *J. Sol. Chem.* **7**, 373–383 (1978).
14. Byrne, R. H. & Kester, D. R. *J. Sol. Chem.* (in the press).
15. Dyrssen, D., Jørgensen, D. & Wengelin, F. *Computer Calculation of Ionic Equilibria and Titration Procedures* (Almqvist & Wiksell, Stockholm, 1968).
16. Byrne, R. H. *Mar. Chem.* **9**, 75–80 (1980).
17. Gažo, J. et al. *Coord. Chem. Rev.* **19**, 253–297 (1976).

The age of the Weddell Basin

John L. LaBrecque

Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA

Peter Barker

Geophysics Section, University of Birmingham, Birmingham B15 2TT, UK

The deep oceanic basin occupied by the Weddell Sea separates the East Antarctic craton from the Antarctic Peninsula, a largely post-Palaeozoic orogenic belt. An understanding of the tectonic evolution of the basin would resolve most of the remaining problems of Gondwanaland reconstruction, which revolve around the relative positions of these two fragments and South America. Norton and Sclater¹ conclude that Indian Ocean marine geophysical data essentially support the Smith and Hallam fit², but the unacceptable overlap of undoubtedly old continental crust^{3–5} on the Antarctic Peninsula and Falkland Plateau implies relative motion between the Peninsula and East Antarctica. Such motion is not required by the reconstruction of Barron and others^{7,8} which, however, can be criticized on other grounds^{9,10}. The marine geophysical observations within the intervening ocean basins have proved useful in formulating plate tectonic reconstructions. Fracture zone trends, seafloor spreading type magnetic anomalies, crustal heat flux, and crustal depth all provide clues to the timing and direction of the continental motions. The acquisition of geophysical data from the Weddell Sea region has been slow^{11–13}. However, the combined interpretation of the two largest data sets extant^{12,13} now being undertaken, seems likely to provide a more tightly constrained Gondwanaland solution. As a first step, we report here an estimated age for the Weddell Sea floor in the area bounded by 60° and 73° S, 5° and 35° W shown in Fig. 1.

Figure 1 also shows the tracks of cruises 1277 and 1578 of ARA Islas Orcadas in the Weddell Basin. Line drawings of seismic reflection profiles obtained on three long north-south lines during these cruises form Fig. 2. A zone W at 64°–65° S¹² separates a more northern area of much greater basement relief from the more southern area being considered here. In this northern area, fracture zone orientations are highly oblique to the gross orientation of the magnetic lineations^{12,13} but the more subdued basement topography to the south suggests that fracture zones there are subdued and have only small offsets. The general depth of the basement increases southwards and, when corrected for sediment loading, is consistent with a Mesozoic age^{12–14} in the area south of zone W with younger crust to the north.

Figure 3 shows our age model for the oceanic crust which floors the Weddell basin. The magnetic anomaly profiles plotted in Fig. 3 are from ref. 13 and are augmented by data collected aboard ARA Islas Orcadas and Project Magnet Flight 720. Note that the magnetic anomalies display a strong east-west lineation and seem to converge to the west. They are compared in Fig. 3 with a magnetic model covering the time period 162–35 Myr. (anomalies M29–13^{15–17}). Seafloor spreading rates vary from 0.6 cm yr⁻¹ during the early Cretaceous to 0.9 cm yr⁻¹ during the Palaeogene. To model the smoothing effect of finite polarity transition zones within the oceanic crust we have convolved the magnetic polarity model with a gaussian function ($\sigma = 3$ km)¹⁸. Rather than assume a magnetic pole position for the anomaly sequence, we have assigned a value of zero skewness to the pattern¹⁹. This effectively defines a family of probable pole positions though we must also account for anomalous skewness²⁰. The choice of zero skewness is based largely on the apparent symmetric nature of anomalies 33–34. All other anomalies can accommodate variations of up to $\pm 30^\circ$ in the model skewness.

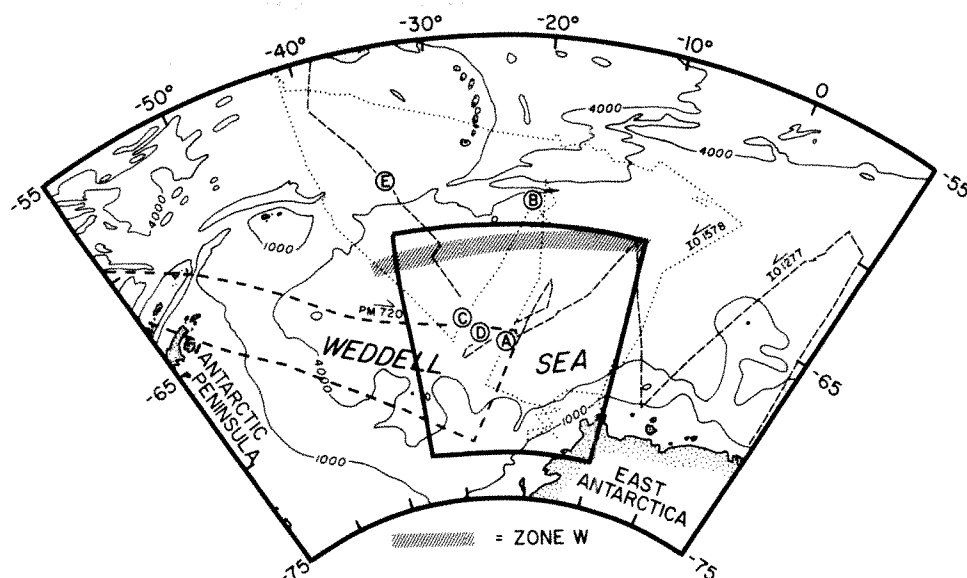


Fig. 1 Map of Weddell Basin showing regional location of Fig. 3. Tracks of *Islas Orcadas* cruises 1578 and 1277 are shown superimposed upon the regional bathymetry in meters²⁹. Encircled letters illustrate locations of seismic reflection profiles of Fig. 2.

Figure 3 also shows synthetic fracture zones or flow lines (stippled lines) calculated for the various reconstructions of Norton and Sclater¹. The flow lines display the displacement of a point on the Antarctica–South America spreading centre with respect to Antarctica fixed²¹. The flow lines are small circles about stage poles calculated for South America motion with respect to Antarctica fixed. We have included flow lines for an M13 reconstruction assuming that the change in plate motions required by the initiation of seafloor spreading in the South Atlantic^{22,23} was accommodated by a change in Antarctica–South America relative motions.

The spreading rates and directions inferred from the calculated flow lines match the observed data reasonably well with two exceptions. The flow line which represents crustal generation during the interval M1–34 (Cretaceous Normal) was reduced in length by 80% from that predicted by the Norton and Sclater poles. This suggests either a significant error in the

Norton and Sclater reconstructions for the time period, or an asymmetry in the seafloor spreading process such as a southward jump of the spreading centre. The dramatic change in spreading direction required by the opening of the South Atlantic in the Valanginian is not observed in the magnetic lineation pattern though a gradual reorientation of the anomaly pattern is perceived.

The strong lineation shown by anomalies 33–34 is consistent in direction and spreading rate with that predicted for Antarctica–South America motion at the Agulhas Magnetic Bight²⁴. The Weddell anomaly 33–34 sequence is a westward extension of the same anomaly sequence observed by Bergh and Barrett²⁵ north-west of the Maud Rise. The calculated lineation direction of the Bergh and Barrett sequence is in error, however, because of their assumption of symmetric seafloor spreading within the Agulhas Basin at the Africa–South America (or Malvinas) spreading centre²⁴.

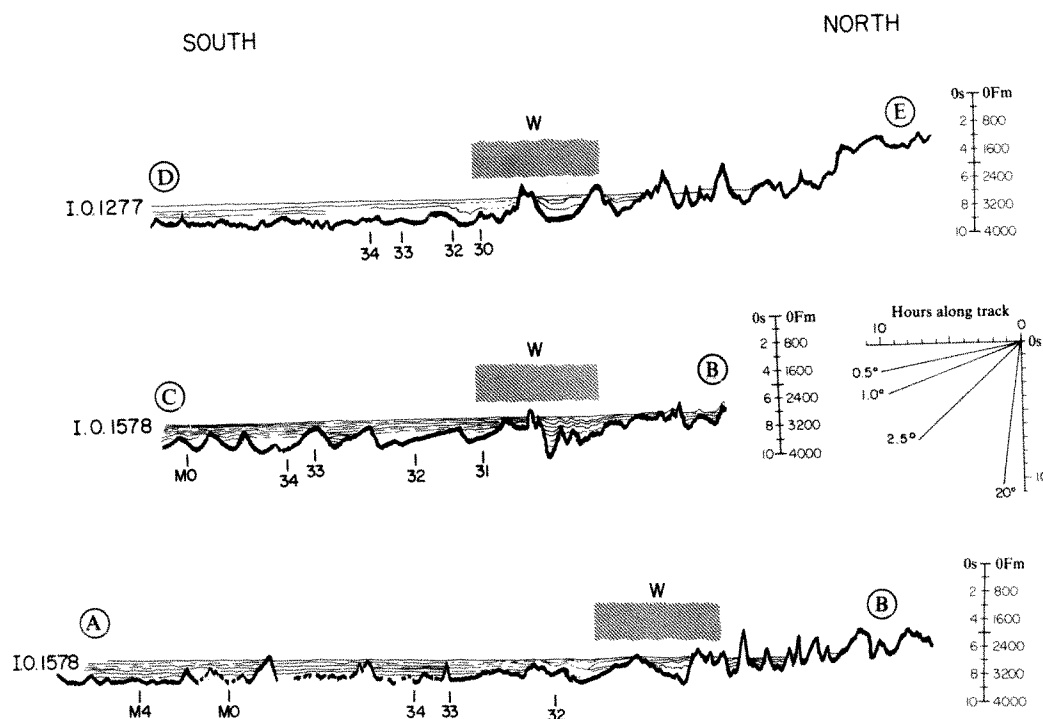
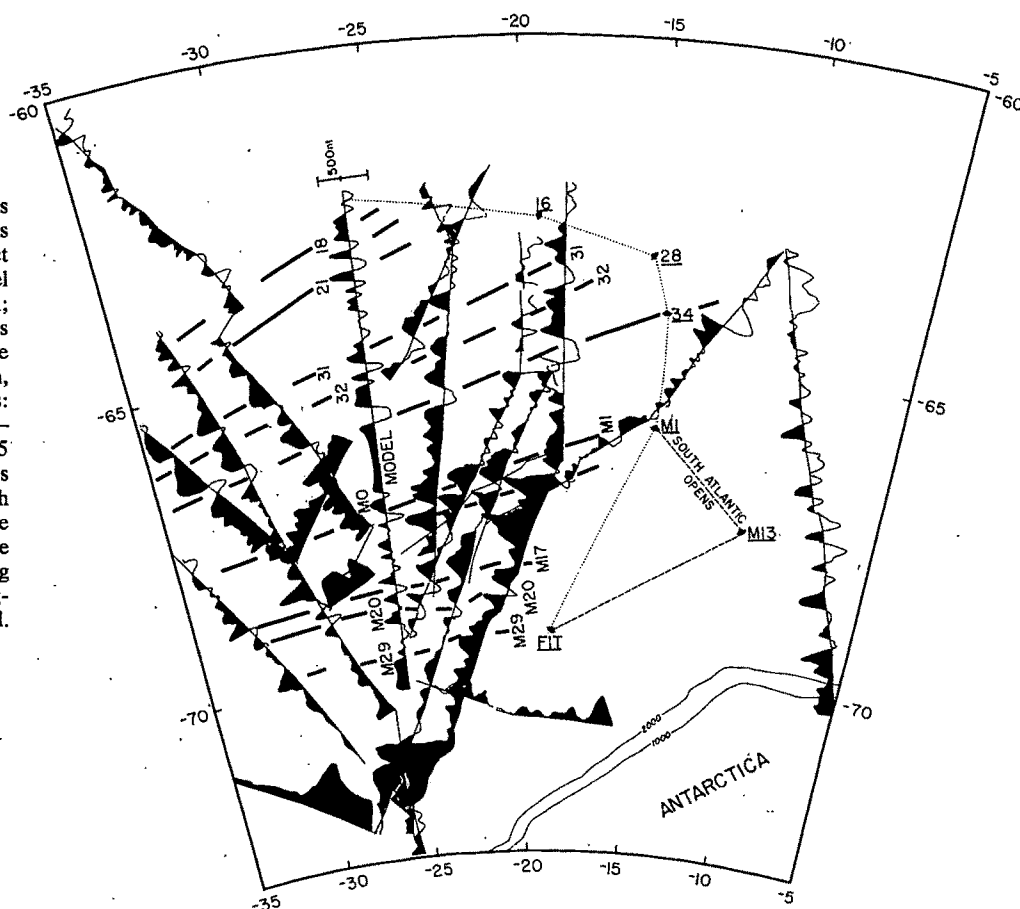


Fig. 2 Line drawings taken from *Islas Orcadas* seismic reflection profiles located in Fig. 1. The locations of identified magnetic anomalies are indicated below each profile. Station numbers are from ref. 26.

Fig. 3 Magnetic anomaly profiles in the Weddell Sea from Islas Orcadas cruises and Project Magnet flight 720. Model parameters: layer depth, 6 km; layer thickness, 500 m, skewness parameter $\theta=0$; transition zone width, $w=12$ km; magnetization, 0.01 emu cm^{-3} , spreading rates: 35–80 Myr BP = 0.9 cm yr^{-1} , 80–108 Myr BP = 45 cm yr^{-1} , 108–165 Myr BP = 0.6 cm yr^{-1} . Flow lines for South American motion with respect to Antarctic fixed are shown as stippled lines and are explained in the text. Ages along flow lines are expressed as magnetic anomalies and are underlined.



Heat flow measurements taken on board Islas Orcadas cruise 1578 reveal a general northerly increase in heat flux within the Weddell Basin²⁶. High heat flow values observed within the Weddell Basin have led to proposals that the Weddell Basin may have formed as a marginal basin during Tertiary times²⁶. The data presented above strongly suggest a much older age for the Weddell Basin. In Figure 4 ages are assigned to the Weddell heat flow stations using the age model presented in Fig. 3, and the Weddell heat flow stations are compared with the global mean values compiled by Parsons and Sclater¹⁴. The high values observed at stations 55 and 40 have been discussed by Zlotnicki *et al.* and may be accounted for by local basement peaks (station 55) and the general variability of individual heat flow values.

The data and model presented above is a first step in tracing the separation of South America and Antarctica since the

breakup of Gondwanaland. Our interpretation of the geophysical data including fracture zone trends, depth to basement, magnetic anomaly lineations, and heat flux indicates that the southern Weddell Basin is pre-late Jurassic to Cretaceous in age. Basement age becomes progressively younger to the north and is Cenozoic in age north of Zone W. We conclude that the Weddell Basin in the area studied was formed at the southern flank of a spreading centre. The northern flank has probably been subducted beneath the Scotia Sea^{12,27,28} or parts may be trapped within it⁶. The general conformity of the observed spreading pattern with that predicted along several lines of evidence for South America–Antarctic relative motion suggests that the spreading centre formed the South American–Antarctic plate boundary throughout much of the breakup of Gondwanaland. The westward convergence of the Weddell magnetic lineation

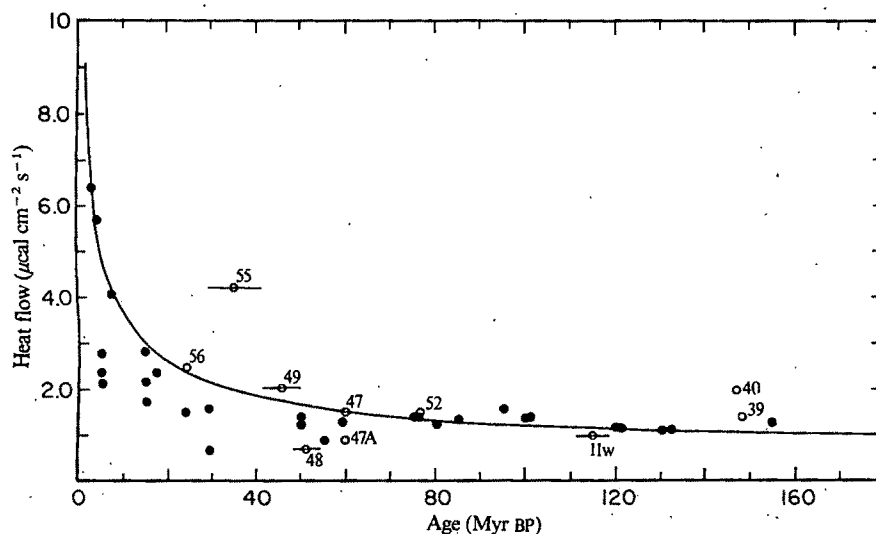


Fig. 4 Comparison of Weddell Basin heat flow values (○) with the Parsons and Sclater heat flow curve (●) using the proposed basement ages.

pattern suggests that the poles of rotation which describe the separation of Antarctic–South America are located very near the Antarctic Peninsula.

This work was supported by NSF grant No. DPP 77-15586 from the Division of Polar Programs.

Received 29 August; accepted 10 October 1980.

1. Norton, I. O. & Sclater, J. G. *J. geophys. Res.* **84**, 6803–6830 (1979).
2. Smith, A. G. & Hallam, A. *Nature* **225**, 139–144 (1970).
3. Adie, R. J. in *Antarctic Research*, 118–162 (Butterworth, London, 1964).
4. Dalziel, I. W. D. in *Antarctic Geoscience* (University of Wisconsin Press, Madison, in the press).
5. Barker, P. F. *et al. Init. Rep. DSDP Let.* **36**, 993–1014 (1977).
6. De Wit, M. J. *Tectonophysics* **37**, 53–81 (1977).
7. Barron, E. J., Harrison, C. G. A. & Hay, W. W. *EOS* **59**, 436–449 (1978).
8. Harrison, C. G. A., Barron, E. J. & Hay, W. W. *Geology* **7**, 374–8 (1979).
9. Dalziel, I. W. D. *Geology* **8**, 260–1 (1980).
10. Godoy, E. & Mpodozis, C. *Geology* **8**, 261–2 (1980).
11. Bregman, M. L. & Frakes, L. A. *Earth planet. Sci. Lett.* **9**, 322–6 (1970).
12. LaBrecque, J. L. & Keller, M. in *Antarctic Geoscience* (University of Wisconsin Press, Madison, in the press).
13. Barker, P. F. & Jahn, R. A. *Geophys. J. R. astr. Soc.* **63**, 271–283 (1980).
14. Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803–827 (1977).
15. Larson, R. L. & Hilde, T. W. C. *J. geophys. Res.* **80**, 2586–2594 (1975).
16. LaBrecque, J. L., Kent, D. V. & Cande, S. C. *Geology* **5**, 330–5 (1977).
17. Cande, S. C., Larson, R. L. & LaBrecque, J. L. *Earth Planet. Sci. Lett.* **41**, 434–440 (1978).
18. Blakely, R. J. *J. geophys. Res.* **79**, 2979–2985 (1974).
19. Schouten, H. & McCamy, K. J. *Geophys. Res.* **77**, 7089–7099 (1972).
20. Cande, S. C. *Earth planet. Sci. Lett.* **40**, 275–286 (1978).
21. Pitman, W. C. III & Talwani, M. *Bull. geol. Soc. Am.* **83**, 619–642 (1972).
22. Larson, R. L. & Ladd, J. W. *Nature* **246**, 209–212 (1973).
23. Rabinowitz, P. D. & LaBrecque, J. L. *J. geophys. Res.* **84**, 5973–6002 (1979).
24. LaBrecque, J. L. & Hayes, D. E. *Earth planet. Sci. Lett.* **45**, 411–428 (1979).
25. Bergh, H. W. & Barrett, D. M. *Nature* **287**, 691–695 (1980).
26. Zlotnicki, V., Sclater, J. G., Norton, I. O. & Von Herzen, R. P. *Geophys. Res. Lett.* **7**, 421–424 (1980).
27. Barker, P. F., Hill, I. A., Weaver, S. D. & Pankhurst, R. J. in *Antarctic Geoscience* (University of Wisconsin Press, Madison, in the press).
28. Hill, I. A. & Barker, P. F. *Geophys. J. R. astr. Soc.* **63**, 427–440 (1980).
29. Heezen, B. C., Tharp, M. & Bentley, C. R. *Antarctic Map Folio Series, Folio 16*, (American Geographical Society, Washington, 1972).

Transcurrent faulting and the anomalous position of pre-Carboniferous Anglesey

M. J. C. Nutt & E. G. Smith

Institute of Geological Sciences (NERC), Ring Road Halton, Leeds LS15 8TQ, UK

The marked differences of pre-Carboniferous geology between Anglesey and of mainland North Wales are accounted for by Devonian transcurrent faulting probably caused by stresses resulting from the closure of the Iapetus Ocean. Displacement along this fault is thought to have exceeded 100 km, so that the Lake District and not Anglesey would formerly have been in juxtaposition with North Wales.

The Carboniferous Limestone of Anglesey can be closely correlated with that of the North Wales mainland¹, so that in Viséan times the two areas probably had the same geographical relationship as today with continuity of sedimentation between them. There is no reason to suppose that the underlying Carboniferous Basement Beds of the two areas do not have a similar relationship. This is in accord with the concept of a Lower Carboniferous sea encroaching southwards from the Irish Sea Basin onto a landmass of older rocks².

These older rocks present a very different picture. The Precambrian of Anglesey consists of a bedded series of arenaceous and argillaceous sediments with subordinate basic volcanics—the Mona complex^{3,4}—ranging in metamorphic facies from greenschist to amphibolite. The only comparable rocks on the mainland are those of southwestern Llyn⁵ and of the River's Mouth Formation at Caernarvon^{3,6}. The Precambrian of the mainland—the Arvonian—consists essentially of acid and intermediate volcanics (of the Bangor and Padarn ridges⁶ and of Brynteg Borehole⁷ in the Harlech Dome), which have suffered only a low degree of metamorphism. Unless these rocks are represented by the Baron Hill Beds and the Bwlch Gwyn

Felsite^{8,9}—an unlikely correlation—the Arvonian seems to be absent in Anglesey.

Greenly³ provisionally assigned several small outliers in Anglesey, including those of the Careg-onen Beds and the Trefdraeth Conglomerate, to the Cambrian, but later¹⁰ considered them to be older. The only fossils recorded in Anglesey that may be of Cambrian age are stromatolites^{11,12} and microfossils¹³ from the Gwna Group, supposed by Greenly³ to be Precambrian rocks. The rocks containing the stromatolites, however, are associated with mélangé and therefore may be derived blocks in much younger deposits. Indeed, there is stratigraphical evidence (unpublished) showing that at least part of the Gwna Mélangé is of Ordovician age. Whether Cambrian rocks are absent in Anglesey or whether they are represented in the Mona Complex (as the above evidence and the stratigraphical interpretation of Barber and Max¹⁴ suggest), there is a marked contrast with the mainland where, only a few kilometres away across the Menai Strait, are some 2,000–3,000 m of slates and grits. Moreover, these latter rocks apparently overlie the Arvonian conformably or with no great unconformity.

The Ordovician rocks of Anglesey, ranging up to basal Caradoc¹⁵, differ from those of the mainland in several major respects. The generally black mudstones of the former contrast with the predominantly grey silty mudstones of the latter, and there is a greater development of arenaceous rocks in Anglesey. Volcanic rocks, which form such a significant part of the mainland Ordovician sequence, are unknown in Anglesey with the probable exceptions of the Fydyln Felsite¹⁴ and the relatively thin lavas and tufts underlying the Middle Llandovery sediments at Parys Mountain¹⁶. Mélanges, a feature of the Ordovician of Anglesey, where they range in age from Arenig to Caradoc, have not been recorded elsewhere in North Wales. In contrast to the mainland, there is evidence of a major pre-Caradoc unconformity in Anglesey¹⁵. This may extend throughout the island, but is best demonstrated in the north, where Caradoc sediments apparently rest on Llanvirn, Arenig and Monian rocks. Upper Caradoc and Ashgill rocks are absent in Anglesey, with the possible exception of the thin volcanics separating Llandovery (*M. argenteus* Zone) and Llanvirn (*D. bifidus* Zone) shales at Parys Mountain. The Ordovician shelly faunas of Anglesey, like those of south-east Ireland and Ramsey Island, Pembrokeshire, have Celtic (Baltic) affinities, whereas those of mainland Wales belong to the Anglo-French (Anglo-Welsh) province^{17–20}.

The only known Silurian rocks in Anglesey are ~50 m of shales at Parys Mountain, attributable to three graptolite zones of the Llandovery. Unlike the equivalent rocks on the mainland, these contain bands of lithic tuff¹⁶. Assuming that the volcanics underlying the Llandovery rocks at Parys Mountain are of Ordovician age, the lower part of the Llandovery sequence of the mainland, which is conformable on the Ordovician, is absent in Anglesey.

The Old Red Sandstone is absent in North Wales except in Anglesey, where it is represented by some 500 m of conglomerates, sandstones and siltstones, probably belonging to the southern British development²¹.

The structure of the Lower Palaeozoic rocks of Anglesey is clearly related in both form and trend to that of the basement. There is evidence that Ordovician sedimentation was controlled by renewed movement on structures within the Mona Complex¹⁷. Although the basement has not been proved on the mainland, it seems most unlikely that the structure of the Lower Palaeozoic rocks reflects that existing in the basement at depth.

This account of fundamental geological differences between Anglesey and the mainland points out the anomalous position of the island with respect to the rest of North Wales in pre-Carboniferous times, suggesting that the two areas did not come into juxtaposition before the late Devonian. So far as the Precambrian is concerned, separation into an oceanic and a continental plate can be effected by a subduction zone—along the Penmynydd Zone of Metamorphism in southeastern Anglesey. Activity on such a zone can perhaps be extended to Tremadoc times—to the end of the Avalonian orogenic

event^{7,22}. During the Ordovician and Silurian periods, however, the operative subduction zone must have lain much farther to the north-west, beyond Anglesey and the Lake District. Closure of the Iapetus Ocean is invoked to account for faunal contrasts between the Ordovician Scoto-Appalachian and Celtic provinces, but the differences between Anglesey and mainland North Wales, belonging respectively to the Celtic and Anglo-French provinces, have not been satisfactorily explained²⁰.

Plate tectonics alone, therefore, do not seem to provide a solution to Anglesey's anomalous position, and it is suggested here that the answer might be found in transcurrent faulting. A northeasterly trending dextral fault with a lateral movement of at least 100 km is envisaged in the vicinity of the Menai Strait (Fig. 1). Such a fault would have a greater displacement than the parallel sinistral Great Glen Fault, but be of lesser magnitude than the subparallel dextral fault postulated²³ along the Iapetus Suture, across Ireland and through the Solway. It has been suggested²⁴, however, that the Great Glen Fault may have undergone up to 100 km of post ~400 Myr dextral movement, and there is doubt²⁵ about the magnitude of the movement along the Iapetus Suture.

The credibility of this speculative transcurrent fault between Anglesey and the rest of North Wales depends on finding areas of compatible geology that could have been adjacent to each before the fracture occurred. In the case of the North Wales mainland the appropriate area is the Lake District.

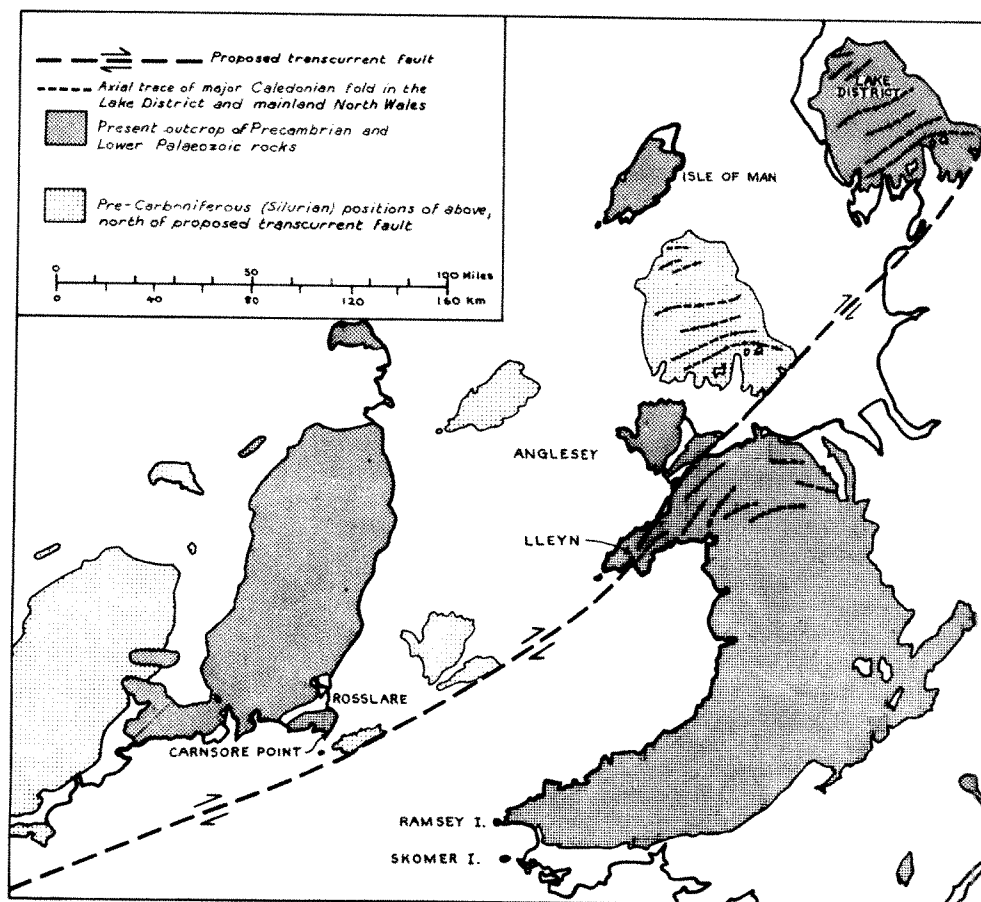
The oldest known rocks in the Lake District are pelitic and arenaceous sediments of the Skiddaw Group, ranging in age from Arenig to Llanvirn. These are comparable to sediments of similar age in North Wales and, like them, contain volcanics—the mid-Arenig tuffs and lavas of the southwestern Lake District²⁶, and the Llanvirn Eycott Volcanics of the northern Lake District²⁷. The Eycott Volcanics are interpreted as the tholeiitic belt of an island-arc environment, and the overlying Llandeilo to Caradoc Borrowdale Volcanics as the calc-alkaline belt²⁸. The Ordovician volcanics of North Wales have been interpreted as part of the calc-alkaline belt in this setting²⁸. Southwards and later this belt gives way to alkaline volcanics. The proposed

transcurrent fault, bringing the two calc-alkaline areas closer together and reducing the width of the belt, as well as eliminating Anglesey from the picture, is advantageous to this model.

Apart from the locally deposited Coniston Limestone and associated beds at the base, the rest of the Lower Palaeozoic sequence in the Lake District is closely comparable with that of mainland North Wales, and it is not unreasonable to suppose that the two sequences were deposited in juxtaposition. The Ashgill Shale Formation²⁹ of the Lake District is lithologically similar to the Bodeidda-Deganwy Mudstones of the Conway area³⁰, the Rebecca Grit of the former³¹ correlating with the Conway Castle Grit of North Wales. The two Llandovery sequences are virtually identical, the Skelgill Beds of the Lake District equating with the succession below the Pale Slates of the Conway area, and the Browgill Beds with the Pale Slates themselves². In the Wenlock there is a marked lithological similarity between the turbiditic Brathay Flags and the Lower Nantglyn Flags, the Lower Coldwell Beds of the Lake District and the Denbigh Grits Group of Wales being proximal facies derived at different times from the west. A prominent band of calcareous concretions in the *Cyrtograptus centrifugus* Zone near the base of the Denbigh Grits Group² compares with a similar band in the same zone in the Lake District and Howgill Fells²⁹. The Middle Coldwell Beds are lithologically identical to the Mottled Mudstones of the Denbigh area, the upper leaf of the former²⁹ having the same zonal position (*Monograptus ludensis*) as the Upper Mottled Mudstone. In the Ludlow the Upper Coldwell Beds equate with the Upper Nantglyn Flags, and the overlying Coniston Grits and Bannisdale Slates are the recognizable equivalents of the lower and upper parts of the Elwy Group respectively. The Coniston Grits are derived dominantly from the north-west²⁹, whilst the slump sheets with sandstones at the same horizon in Wales come from the west². Higher Silurian horizons exposed in the Lake District are not known in North Wales.

The trends of the Caledonian folding in the Lake District and North Wales, swinging from SW-NE to W-E towards the east, show a better fit when the two areas are in north-south

Fig. 1 Map showing the proposed transcurrent fault and the pre-Carboniferous (Silurian) positions of Anglesey and the Lake District relative to Wales. The line of the fault approximates in the north-east to the Dent Line, passing southwestwards beneath Carboniferous and later cover to run along the Menai Strait (the Dinorwic Fault), where transcurrent movement has previously been deduced¹⁴ from the extensive development of mylonites. The fault crosses the Lleyn (where it may have been dislocated by subsequent transcurrent faulting) from Porth Nefyn to the western side of Hell's Mouth, confining the Precambrian rocks, which have affinities with Anglesey, to its north-western side and separating substantially different Ordovician sequences^{5,36-38}.



juxtaposition than in their present geographical positions (Fig. 1).

If the Lake District and the North Wales mainland are matching parts of an Ordovician island-arc sequence and of a Silurian basin of sedimentation torn apart by the proposed transcurrent faulting, we must look southwestwards for the area originally adjacent to Anglesey, if it still exists. The obvious candidate is south-west Wales (Pembrokeshire and Carmarthenshire). Here, although the Cambrian is markedly unconformable on the Precambrian and there are Ordovician faunal similarities between Ramsey Island and Anglesey as well as agreement between the heavy mineral assemblages of the Old Red Sandstone of Pembrokeshire and Anglesey²¹, the general affinities of south-west Wales are with the North Wales mainland. Thus the Uriconian-type volcanic sequence composing the Precambrian of Pembrokeshire compares with the Arvonian rather than the Mona Complex, there is a thick Cambrian succession, an Ordovician sequence containing limestones, thick volcanics and no recorded mélanges, and, in west Pembrokeshire, a well developed Silurian succession extending through the Wenlock and Ludlow into the Downtonian.

Anglesey is more comparable with southeastern Ireland. The Precambrian para- and ortho-gneisses with mylonites of the island and of the Lleyn Peninsula¹⁴ are not dissimilar from those of the Rosslare Complex³². The basic intrusions forming the Rosslare dyke swarms³³ may be represented in Anglesey by basic dykes cutting gneisses, for example, at Treferwyd¹⁴. The Coedana Granite of Anglesey and the Sarn Granite of the Lleyn belong to the same orogenic episode as the Carnsore Granite intruded into the Rosslare Complex³⁴. The New Harbour Group of Anglesey is similar to the ?Precambrian Cullenstown Formation¹⁴ and the South Stack Formation to the Lower to Middle Cambrian Bray Group³. As in Anglesey, there is a pre-Caradoc unconformity in southeastern Ireland, though the overlying rocks of the latter, the Duncannon Group, are predominantly volcanic³⁵. There is insufficient detailed knowledge of the Llandovery rocks of southeastern Ireland to attempt a comparison with Anglesey. Structural trends (east-northeasterly to northeasterly) in this part of Ireland are similar to those in Anglesey. The form of the folds in the Irish Lower Palaeozoic is also reminiscent of the situation in the Monian and Ordovician of Anglesey.

The line of the proposed transcurrent fault is shown in Fig. 1.

South-west of the Lleyn the fault probably runs down St George's Channel to lie off the coast of south-east Ireland (Fig. 1), but possibly swings farther west through Rosslare Harbour and along the mylonite zone of the Rosslare Complex^{14,39,40}, separating the gneisses of the latter from the rest of Ireland. An offshore position is favoured, however, because of geophysical evidence of a major fault in this region^{41,42} and the likelihood that the northern boundary of the subsequent Celtic Sea Basin⁴³ corresponds to the line of such a major pre-existing fault^{39,40,44}.

A dextral movement of at least 100 km is envisaged along the proposed fault (Fig. 1). This movement not only removes Anglesey from its present incongruous proximity to North Wales, but creates a more rational relationship between the latter area and the Lake District. A more southwesterly location for Anglesey would also bring the Silurian tuffs of the island closer to the zone of known Silurian volcanism extending from the Dingle Peninsula, through Skomer, to the Mendips⁴⁵. Similarly a more southerly position for the Lake District diminishes the inclination of a southeasterly dipping Lower Palaeozoic Benioff Zone originating along a trench approximately through the Solway. K_2O/SiO_2 plots for Borrowdale volcanic rocks (unpublished data) suggest, following studies of potash-depth relations⁴⁶, that the depth of the Benioff Zone below the central Lake District was of the order of 235 km. This gives a dip from the Solway (allowing for 50% crustal shortening⁴⁷) of $\sim 65^\circ$, an excessive figure in comparison with that quoted⁴⁶ for modern continental margin arcs ($\sim 40^\circ$). The re-location of the Lake District as in Fig. 1 gives a Benioff Zone dip

of $\sim 45^\circ$, which angle would be further reduced by a more southwesterly Ordovician location.

The cause of the dextral transcurrent movement can be looked for in NW-SE compressional stresses following the final closure of the Iapetus Ocean in Devonian times, the same explanation invoked²³ for similar lateral movement on the line of the Iapetus Suture itself. Major transcurrent movements affecting the Caledonides of the British Isles and North America in the mid-Devonian and extending over ~ 20 Myr are indicated by palaeomagnetic studies⁴⁸.

The Devonian movement may have been merely the culmination of dextral wrenching initiated in the Cambrian or Precambrian. Whilst Fig. 1 shows the best fit of the Lake District and North Wales for the Silurian rocks, there is evidence that a more southwesterly position for the Lake District would be better for a match of the Ordovician rocks of the two areas. The fault might, therefore, date back to the Cambrian or Precambrian and its line might have been dictated by the position of the Penmynydd subduction zone during its active lifetime. This situation would be similar to, for example, the present position in New Zealand, where, at an active plate junction, there is an intimate relationship between the subduction zone forming the Kermadec-Tonga Trench and the subparallel strike-slip Alpine Fault⁴⁹.

It is feasible that major northerly and northwesterly faults in North Wales and northwestern England originally developed as conjugate fractures to the proposed transcurrent faulting. The age (around 355 Myr as shown by K-Ar dating) and direction (ENE-WSW) of the main suites of metalliferous veins in Anglesey (Parys Mountain¹⁶), the Lake District (Vale of Newlands⁵⁰), Central Wales (Plynlimon mining field⁵¹) and Salop (Shelve area⁵¹) suggest that these veins may occupy tensional structures associated with dextral transcurrent faulting of the proposed trend, a situation in keeping with the theory⁵² of hydraulic fracturing.

This paper is published with the permission of the Director, Institute of Geological Sciences (NERC).

- Power, G. & Somerville, I. D. *Proc. Yorks. geol. Soc.* **40**, 491-497 (1975).
- Warren, P. T., Price, D., Nutt, M. J. C. & Smith, E. G. *Mem. geol. Surv. G.B.* (in the press).
- Greenly, E. *Mem. geol. Surv. G.B.* (1919).
- Shackleton, R. M. in *The Pre-Cambrian and Lower Palaeozoic Rocks of Wales* (ed. Wood, A.) 1-22 (University of Wales, 1969).
- Matley, C. A. Q. *Jl. geol. Soc. Lond.* **84**, 440-504 (1928).
- Greenly, E. Q. *Jl. geol. Soc. Lond.* **100**, 269-287 (1944).
- Allen, P. M. & Jackson, A. A. *Bull. geol. Surv. G.B.* No. 61 (1968).
- Wood, D. S. in *The Pre-Cambrian and Lower Palaeozoic Rocks of Wales* (ed. Wood, A.) 47-66 (University of Wales, 1969).
- Wood, D. S. *Soc. econ. Paleont. Miner. Spec. Publ.* No. 19, 334-344 (1974).
- Greenly, E. *Geol. Mag.* **83**, 237-240 (1946).
- Wood, M. & Nicholls, G. D. *Nature phys. Sci.* **241**, 65 (1973).
- Downie, C. *Geol. Soc. Lond. Spec. Rep.* No. 6, 113-115 (1975).
- Muir, M. D., Bliss, G. M., Grant, P. R. & Fisher, M. J. *J. geol. Soc. Lond.* **136**, 61-64 (1979).
- Barber, A. J. & Max, M. D. *J. geol. Soc. Lond.* **136**, 407-432 (1979).
- Bates, D. E. B. *Geol. J.* **8**, 29-58 (1972).
- Nutt, M. J. C., Ineson, P. R. & Mitchell, J. G. *Geol. Soc. Lond. Spec. Publ.* No. 8, 619-627 (1979).
- Bates, D. E. B. *Bull. Br. Mus. nat. Hist. (Geol.)*, **18**, 1-28 (1969).
- Williams, A. in *The Pre-Cambrian and Lower Palaeozoic Rocks of Wales* (ed. Wood, A.) 117-150 (University of Wales, 1969).
- Williams, A. *Spec. Pap. Palaeont.*, No. 12, 241-269 (1973).
- Neuman, R. B. & Bates, D. E. B. *Palaeontology*, **21**, 571-613 (1978).
- Allen, J. R. L. *Proc. Yorks. geol. Soc.* **35**, 139-185 (1965).
- Rodgers, J. *Am. J. Sci.* **265**, 408-427 (1967).
- Phillips, W. E. A., Stillman, C. J. & Murphy, T. J. *geol. Soc. Lond.* **132**, 579-609 (1976).
- Halliday, A. N., Aftalion, M., van Breemen, O. & Jocelyn, J. *Geol. Soc. Lond. Spec. Publ.* No. 8, 653-661 (1979).
- Johnson, M. R. W., Sanderson, D. J. & Soper, N. J. *Geol. Soc. Lond. Spec. Publ.* No. 8, 165-186 (1979).
- Turner, R. E. & Wadge, A. J. *Proc. Yorks. geol. Soc.* **42**, 405-414 (1979).
- Millward, D., Moseley, F. & Soper, N. J. *Yorks. geol. Soc. occ. Publ.* No. 3, 99-120 (1978).
- Fitton, J. G. & Hughes, D. J. *Earth planet. Sci. Lett.* **8**, 223-228 (1970).
- Ingham, J. K., McNamara, K. J. & Rickards, R. B. *Yorks. geol. Soc. occ. Publ.* No. 3, 121-129 (1978).
- Elles, G. L. Q. *Jl. geol. Soc., Lond.* **65**, 169-194 (1909).
- Rose, W. C. C. & Dunham, K. C. *Econ. Mem. geol. Surv. G.B.* (1977).
- Max, M. D. *Geol. Soc. Lond. Spec. Rep.* No. 6, 97-101 (1975).
- Baker, J. W. *Proc. R. Ir. Acad.* **67B**, 177-184 (1968).
- Wright, A. E. in *North Atlantic Geology and Continental Drift* (ed. Kay, M.) 93-109 (American Association of Petroleum Geologists, 1969).
- Bruck, P. M. et al. *Geol. Soc. Lond. Spec. Publ.* No. 8, 533-544 (1979).
- Matley, C. A. Q. *Jl. geol. Soc. Lond.* **88**, 238-273 (1932).
- Nicholas, T. C. Q. *Jl. geol. Soc. Lond.* **71**, 83-143 (1915).
- Matley, C. A. Q. *Jl. geol. Soc. Lond.* **94**, 555-606 (1938).
- Rast, N. *Geol. Soc. Lond. Spec. Publ.* No. 3, 197-213 (1969).
- Baker, J. W. *Geol. Mag.* **110**, 447-455 (1973).

41. Blundell, D. J., Davey, F. J. & Graves, L. J. *J. geol. Soc. Lond.* **127**, 339–375 (1971).
42. Dobson, M. R., Evans, W. E. & Whittington, R. *Rep. Inst. geol. Sci.* No. 73/11 (1973).
43. Ziegler, P. A. *Proc. Conf. on Petroleum Geology of the Continental Shelf of North West Europe*, London (in the press).
44. Dunning, F. W. & Max, M. D. *Geol. Soc. Lond. Spec. Rep.* No. 6, 11–14 (1975).
45. Stillman, C. J. & Francis, E. H. *Geol. Soc. Lond. Spec. Publ.* No. 8, 557–577 (1979).
46. Dickinson, W. R. *Geology* **3**, 53–56 (1975).
47. Oertel, G. *Bull. geol. Soc. Am.* **81**, 1173–1188 (1970).
48. Morris, W. A. *Can. J. Earth Sci.* **13**, 1236–1243 (1976).
49. Badgley, P. C. *Structural and Tectonic Principles* (Harper and Row, New York, 1965).
50. Ineson, P. R. & Mitchell, J. G. *Geol. Mag.* **111**, 521–537 (1974).
51. Ineson, P. R. & Mitchell, J. G. *Trans. Instn. Min. Metall.* **84**, B7–16 (1975).
52. Phillips, W. J. *J. geol. Soc. Lond.* **128**, 337–359 (1972).

Evidence of syndimentary tectonic movements in the Triassic halite of Cheshire

Roger M. Tucker & Maurice E. Tucker

Department of Geology, The University, Newcastle upon Tyne NE1 7RU, UK

During the Triassic, north-west Europe was subjected to tensional stresses which resulted in the formation of a complex system of rapidly subsiding grabens and wrench-faulted basins. This pattern of regional crustal extension, which is part of the larger scale Mesozoic break-up of the Pangean megacontinent, is related to the Triassic opening of the Tethys Ocean in southern Europe and rifting in the Arctic North Atlantic, and is a prelude to the Jurassic opening of the southern North Atlantic¹. Great thicknesses of generally continental sediments, often with salt deposits, accumulated in the Triassic basins, which were mainly avolcanic. In western Britain, a complex series of fault-bounded basins extended from the Western Approaches and Channel area northwards to the Irish Sea, and includes the Worcester graben and the Cheshire Basin. The present study of the halite formations in the Cheshire Basin shows evidence for pencon-temporaneous fault movements; these were a feature of Triassic times but it is rare that such movements can be detected so precisely as is the case here.

The Cheshire Basin is filled with Triassic sediments dating from the Anisian (lower Middle Triassic) and reaching several kilometres in thickness^{2,3}. There are two thick halite developments in the sequence: a lower Northwich Halite Formation (290 m) and an upper Wilkesley Halite Formation (400 m)⁴. Each of these consists of a series of halite units up to 40 m thick separated from each other by horizons of dolomitic silty clay ('marl'). Between the halite formations, and above and below, there are thick formations of marl of the Mercia Mudstone Group, previously referred to as the Keuper Marl. In the ICI mine at Winsford, Cheshire, where these studies were undertaken, halite is being extracted from two beds in the lower part of the Northwich Halite Formation.



Fig. 1 Photograph of the roof of the mine at Winsford, Cheshire, showing large-scale polygonal pattern (polygon is 14 m wide).

Although principally flat-lying, the halite units in the Cheshire Basin are locally folded into monoclinical structures, which may impart a 45° dip on the beds. The trends of the monoclines are controlled by faults in the underlying Carboniferous basement. Several of the monoclinical structures, trending NE–SW, occur in the area now being mined at Winsford.

The Cheshire halite shows well defined bedding on a scale of 10–50 mm, resulting from an alternation of pure and less pure halite. The salt was deposited from standing bodies of water which were at most only several metres deep⁵. Bromine analysis indicates a marine source for the water (R. M. T., in preparation), although the nearest seawater at the time would have been many hundreds of kilometres distant in southern or northernmost Europe.

On bedding surfaces of the halite, best seen on the roof of the mine, spectacular polygonal patterns are present^{6,7}. These polygons are up to 14 m wide and are defined by 0.5 m-thick bands of secondary salt interlayered with clastic material (Fig. 1). Extending down to 6 m from the polygons, and so seen on mine walls, are tapering V-shaped fissures, with the same layered fills. The layering in the fissures is arranged as a series of

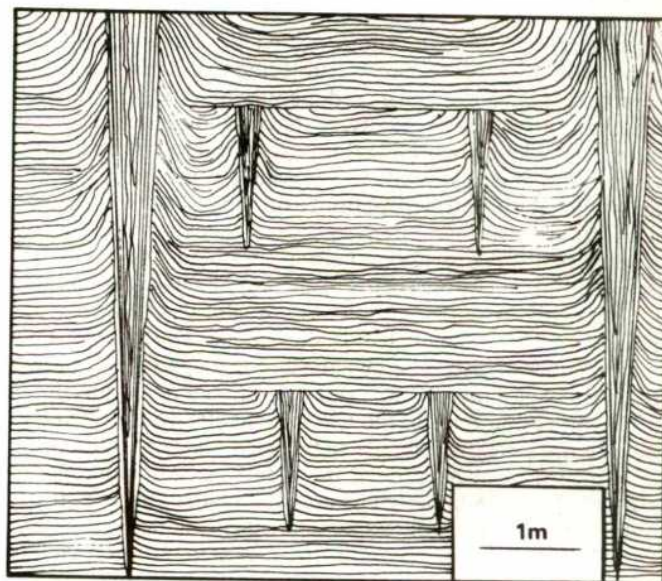


Fig. 2 Sketch from photograph of the wall of the mine at Winsford, Cheshire, showing deep V-shaped fissures and upturned bedding between each fissure.

10 mm wide subvertical bands running parallel to the sides of the fissure. Notably, the fine bands in the fissure fills are paired with a mirror-symmetry about a central unit of secondary salt. Bedding in the halite is upturned adjacent to the fissures (Fig. 2). We propose that each of the clastic layer pairs represents a single opening and filling of the original crack. Capillary brine drawn up the crack, from an underlying ground water table, precipitated the secondary halite. The highly layered nature of the fissure fills thus represents successive openings and fillings of the crack.

The formation of the polygons and fissures is attributed to a process of thermal contraction, acting on the halite after its deposition when it was subaerially exposed⁷. The formation of thermal contraction polygons in halite is very similar to that of periglacial polygons where thermal contraction of ice is involved⁸. Indeed, the thermal properties of salt and ice have much in common, both having high coefficients of linear thermal expansion. When an exposed salt surface is subjected to temperature fluctuations a polygonal pattern, defined by open fissures extending down into the salt bed, is formed during colder periods. The open fissures may then be filled with clastic material transported to the area by wind or water. During subsequent surface temperature increases, the salt surface re-expands and causes compaction of the clastic fissure fill material.

contraction-crack infilling-expansion of the surface salt causes bending up of the salt beds adjacent to the fault. The polygons in Cheshire show that there were long intervals of exposure and non-deposition after evaporation to dryness of the water body from which the halite was precipitated.

It can be demonstrated⁷ that the thermal contraction cracks were induced by temperature fluctuations operating over an annual period. Such temperature fluctuations are well documented from modern deserts, including Death Valley, California, where dry salt lakes with polygonal cracks are known⁹.

In some areas of the mine at Winsford, polygons are not developed, but in their place stripes occur. These form two linear belts, trending NE-SW, which can be traced across the whole mine area. The internal layering of the stripes and their associated fissures are identical to those of the polygons and their fissures, so that a process of thermal contraction was also involved. The location and orientation of the linear stripe patterns are significant: they run parallel to the axes of two monoclinical fold structures and are only developed close to the upthrow side of the monoclines, where they replace the polygons (Fig. 3).

Patterns of stripes associated with polygons have been described from several locations^{10,11}. Neal¹¹ demonstrated that stripes would form in sediments subjected to desiccation if they were located on a slope. The sediments would be pre-stressed, because of the effects of gravity, so that on contraction the stress pattern would be mirrored in the pattern of desiccation fissures produced: that is, a striped pattern which trends normal to the dip of the slope. Striped patterns on slopes associated with polygons are also found in the periglacial environment in ice-rich sediments. A process of thermal contraction (but of ice) is also involved, but where stripes form, they are always orientated parallel to the dip of the slope. The reason for this is that ice-rich sediments have the ability to flow downslope. This has the effect of modifying the original polygon form through several phases from stretched polygons to stripes orientated down the slope.

The striped patterns in the Cheshire halite are thus taken to indicate the development of a slight slope on the surface of the halite which must have been the result of contemporaneous movements on the underlying basement faults. That the stripes in the halite are parallel to the monoclinical fold axes, and thus normal to the dip of the slope that existed, shows that the movements on the faults caused the salt above the fault to become stressed, due to the bed being stretched over the fault. When thermal contraction cracking affected the newly exposed salt surface, polygonal patterns formed on the unstressed areas away from the fault-monoclinical axes and striped patterns following the stress pattern formed on the pre-stressed areas above the basement faults.

The thermal contraction polygons and stripes in the Triassic Cheshire halite show that there were contemporaneous movements on Carboniferous basement faults. If the fault movements occurred after thermal contraction cracking had begun, the

resultant crack pattern close to the upthrow side of the fault would be extremely complex, being made up of both polygons and stripes. However, the striped areas in the Winsford mine have a very simple pattern containing only stripes. This then indicates that the fault movements must have occurred either immediately after halite precipitation had ceased, and the salt bed became emergent, or during salt deposition, and it was this movement that caused the salt bed to become subaerially exposed. It is likely that such movements affected sedimentation in the Cheshire Basin in other ways, such as contributing towards the periodic cutting off of the Cheshire Basin from its distant marine water source, and affecting the siliciclastic supply for the basin by controlling rates of uplift of surrounding highlands.

We thank ICI Mond Division for permitting us to undertake geological studies in the salt mine at Winsford. R.M.T. acknowledges NERC for financial support during this study, and both authors acknowledge the respective Tucker wives.

Received 7 November 1980; accepted 3 February 1981.

1. Ziegler, P. A. 26th Congr. Geol. Int. Colloque C6, 249-280 (1980).
2. Warrington, G. *Q.Jl. geol. Soc. Lond.* **126**, 183-223 (1970).
3. Evans, W. B. *Q.Jl. geol. Soc. Lond.* **126**, 103-123 (1970).
4. Warrington, G. *et al. Geol. Soc. Lond. Spec. Rep.* 13 (1980).
5. Artherton, R. S. *Geol. J.* **15**, 43-58 (1980).
6. Hollingworth, S. E. *Proc. geol. Soc.* **1623**, 93 (1965).
7. Tucker, R. M. *J. sedim. Petrol.* (in the press).
8. Lachenbruch, A. L. *Geol. Soc. Am. Spec. Pap.*, 70 (1962).
9. Hunt, C. B. & Washburn, A. L. *U.S. Geol. Surv. Prof. Pap.* 494B, B104-B133 (1966).
10. Washburn, A. L. *Bull. geol. Soc. Am.* **67**, 823-865 (1956).
11. Neal, J. T., in *Playa Lake Symp. Proc.* (ed. Reeves, C. C. Jr.) 107-132 (ICASALS Publ. 4, Texas Technical University, Lubbock, 1972).

Functional reconstruction of gait from the Pliocene hominid footprints at Laetoli, northern Tanzania

J. Charteris, J. C. Wall & J. W. Nottrodt

School of Human Biology, College of Biological Science, University of Guelph, Guelph, Ontario, Canada N1G 2W1

Direct measurements of fossilized Pliocene hominid footprints yield only spatial information, such as the length and angle of step¹⁻⁴. A more complete description of gait requires temporal information so that parameters such as the number of steps taken per minute (cadence) or speed of walking can be calculated. The relationships between step length, cadence and speed of walking have been worked out for modern man and regression equations of high predictability are available⁵. When measurements from the hominid footprints are fitted to these equations a pattern of gait emerges which contradicts previous reconstructions. As we report here, the pattern of gait consistent with the reported footprint measurements from Laetoli is one that might be described as 'strolling', in which the stride is quite short, even when related to the maker's height.

The data reported from Laetoli include the lengths of each foot (*l*) and 'stride', but contrary to the most widely accepted convention the latter actually represent steps or half-strides. Usually stride-length (*L*) is defined as the distance between two consecutive points of contact by the same foot, which means that a stride is made up of two steps. This distinction between step and stride is critical in attempting to reconstruct gait. White⁴ has assessed the stature (*S*) of the makers of the Laetoli footprints; he used data obtained from modern, small people, the Mawabi Pygmies, whose mean foot-length he found to be equivalent to 15.5% of stature. Independent assessment of stature of the East African hominids agrees that they were very small^{6,7}. Thus, from the measurement of stride-length and the estimated stature we can calculate relative stride-length (*L'*).

$$\text{Relative stride-length, } L' = \frac{\text{Stride-length (m)}}{\text{Stature (m)}}$$

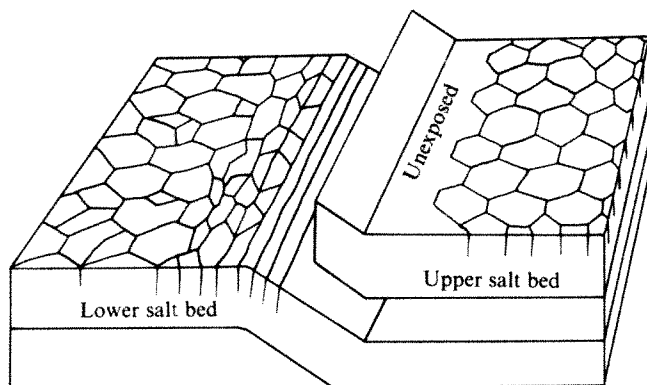


Fig. 3 Schematic block diagram showing the relative positions of polygons and stripes in the ICI mine at Winsford.

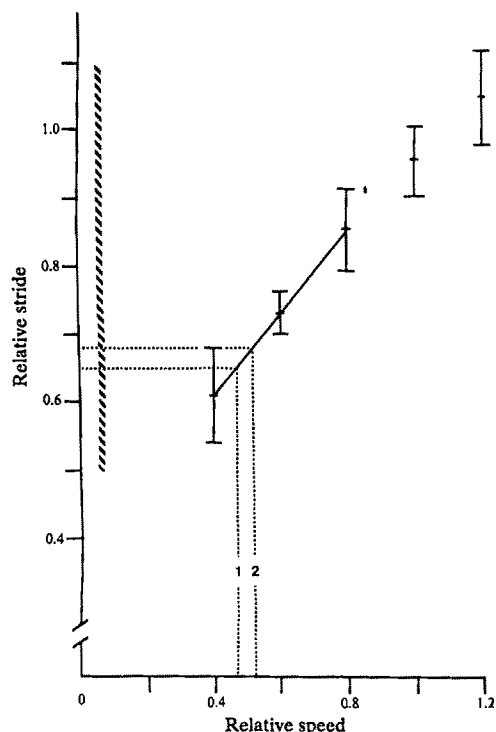


Fig. 1 Graph of stride-length (in units based on fractions of stature) against speed (in units based on fraction of stature covered overground per second) for a normal adult population. In the small portion of the range of relative speeds relevant to this analysis the relationship is linear. The bar on the ordinate reflects the normal range of relative strides available to human adults. The relative speeds of the early hominids whose footprints have been reported indicate a slow traverse across the Laetoli ash beds. Table 1 shows how the ordinate values for hominids leaving trails 1 and 2 are obtained.

Measurements of gait, such as time and length of stride, vary with walking speed. When walking together, at a speed calculated in standard units such as m s^{-1} , a tall and a short subject have very different step-lengths and cadences, the shorter individual having to increase one and/or the other, in order to 'keep up'. These differences in gait due to stature may be factored-out by measuring speeds in units of body length covered overground per second^{5,8,9}; thus:

$$\text{Relative speed, } V' = \frac{\text{Velocity (m s}^{-1}\text{)}}{\text{Stature (m)}}$$

Table 1 Gait parameters measured and derived from the Laetoli footprints

Measure	Units	Symbol	Hominid trails		Source or derivation
			1	2	
Foot-length	m	<i>l</i>	0.185	0.215	Direct measurements
Stride-length	m	<i>L</i>	0.774	0.944	Ref. 2
Foot-length-per stride	—	<i>l'</i>	4.18	4.39	
Stature	m	<i>S</i>	1.19	1.39	$S = 100l/15.5$ Ref. 4
Relative stride	Stat	<i>L'</i>	0.65	0.68	$L' = L/S$
Relative speed	Stat. s^{-1}	<i>V'</i>	0.47	0.52	$L' = [60 V'(1 - \beta_2)]/\alpha_2$ Ref. 5
Velocity	m s^{-1}	<i>V</i>	0.56	0.72	$V = V'S$
Stride time	s	<i>T</i>	1.38	1.31	$T = L/V$
Cadence	Steps min^{-1}	<i>C</i>	87	92	$C = 120/T$

The curvilinear equation for predicting V' from L' is

$$L' = \frac{60 V' (1 - \beta_2)}{\alpha_2} \quad (1)$$

where constants $\alpha_2 = \pm 63$ (males), ± 69 (females), and $\beta_2 = 0.58$.

Figure 1 plots relative speed against relative stride using Grieve's equation and our own data collected from a population of 13 normal young adult caucasian Canadians (6 males, 7 females). Using the measured stride-length and estimated stature, it is possible to calculate values for the relative strides of the early hominids and to use these values to estimate relative speed (Fig. 1).

We use here data from footprints of the two hominids reported by Leakey and Hay² from site G in fossil locality 8 at Laetoli. They assigned the 22 smaller prints to one individual, designating these as trail 1, and the 12 larger prints to a second individual, assigned to trail 2. (Evidently there is far more assurance of these trails than of those found at site A.) The data measured from the footprints at site G, together with the gait parameters derived from them, are shown in Table 1.

Given that modern man normally walks at about 0.86 relative speed^{5,9}, it can be seen from Fig. 1 and Table 1 that the Laetoli hominids walked slowly (around 0.50). Recognizing that in modern man almost every measurable aspect of gait is related to speed, we suggest that in field work the walking speed can be

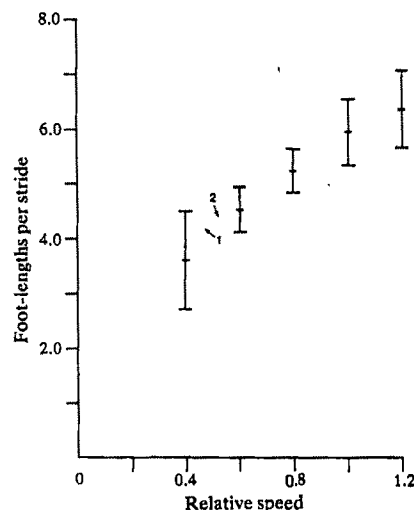


Fig. 2 Graph of stride-lengths in units of foot-length against relative speed, for the same normal young adult human group, showing mean values ± 1 s.d. Based on the relative speed assessment from Fig. 1, the appropriate ordinate values, measured directly off the substrate are plotted, confirming without need of stature estimates of the hominids that their pattern of gait was essentially human.

determined on the basis of the foot-length and the length of stride, which can be measured directly from a set of footprints. Through the most used portion (0.4–1.0) of the range of relative speeds open to modern man there is an approximately linear relationship between the number of foot-lengths in one stride and relative speed (Fig. 2). Thus, even without estimates of stature it is possible to determine whether the gait was slow, medium or fast, and whether footprints made side-by-side are likely to suggest individuals walking along together or traversing the same path but at different speeds and times. Figure 2 confirms our reconstruction of slow speeds, not just slow in human terms, but slow even for these small hominids.

We have not intended to make definitive statements about the characteristic gait of the Laetoli hominids, but rather to show workers in the field how they might reconstruct gait patterns of early bipedal hominids when further trails of footprints are found. Our calculations using data published from Laetoli are based on certain assumptions: that the measurements made on the footprints are accurate, that the prints are indeed those of

early bipedal hominids and that in any sequence of two or three prints only one individual is represented. We used White's estimate of stature⁴, but even using the upper and lower limits from White's tables, and values from groups other than the Mawabi Pygmies our conclusions would not be substantially altered.

If we assume that the relative proportions of the feet of Pliocene hominids are essentially the same as those of modern man⁴, and that their gait was functionally equivalent¹⁰, our determinations of the walking patterns of the hominids (Table 1) are valid, for they are the values that would be expected from modern human subjects of small stature.

Our reconstruction challenges previous claims³ that the stride is "... quite long relative to the creature's small size". On the contrary, it is quite short relative to its maker's stature. Grieve¹¹ has given the adult range of relative strides (Fig. 1) from which it is apparent that a relative stride-length ~0.65 is closer to the lower end of the normal range. Normal young adults select 'comfortable' or 'preferred' walking speeds remarkably close to the 0.86 relative speed^{9,12,13}. Grieve^{5,11,14} has pointed out that walking speeds are situation dependent, but we believe that unless some environmental or motivational constraint was acting on these hominids to walk slowly (in the sense of cautiously), the patterns thus far reported from the ashfall tuff of Laetoli are patterns consistent with what humans call 'strolling'. In these conditions there is no need to invoke a narrow step-width or variable step-angle as evidence of a pattern in any way different from that of modern man; normal humans, strolling at around 0.4 relative speed, often exhibit the same features.

Received 5 December 1980; accepted 11 February 1981.

1. Leakey, M. D. *Antiquity* **52**, 133 (1978).
2. Leakey, M. D. & Hay, R. L. *Nature* **278**, 317-323 (1979).
3. Leakey, M. D. *Natn. Geogr.* **155**, 446-457 (April 1979).
4. White, T. D. *Science* **208**, 175-176 (1980).
5. Grieve, D. W. & Gear, R. J. *Ergonomics* **5**, 379-399 (1966).
6. Johanson, D. C. & Taieb, M. *Nature* **260**, 293-297 (1976).
7. Johanson, D. C. & White, T. D. *Science* **203**, 321-330 (1979).
8. Charteris, J., Leach, D. & Taves, C. J. *Anat.* **128**, 803-819 (1979).
9. Charteris, J. & Taves, C. J. *Perception Motor Skills* **47**, 659-666 (1978).
10. Lovejoy, C. O., Heiple, K. G. & Burnstein, A. H. *Am. J. Phys. Anthropol.* **38**, 757-780 (1973).
11. Grieve, D. W. *Biomed. Engng* **3**, 119 (1968).
12. Wall, J. C. & Charteris, J. *Ergonomics* **23**, 425-435 (1980).
13. Murray, M. P., Drought, A. B. & Kory, R. C. *J. Bone Jt Surg.* **A46**, 335-360 (1964).
14. Grieve, D. W. *Physiotherapy* **55**, 452-460 (1969).

Firefly mate-rivals mimic their predators and vice versa

James E. Lloyd

Department of Entomology and Nematology, University of Florida, Gainesville, Florida 32611, USA

Fireflies, beetles of the family Lampyridae, use flashes of bioluminescent light for sexual communication—by means of coded patterns, males and females identify and locate each other in darkness¹. In sexual signalling of *Photinus macdermotti* (eastern United States) the male emits two flashes separated by ~2 s and the female then waits ~1 s before replying with a single flash (Figs 1a, 2a)². Females of *Photuris versicolor* and *Photuris 'B'* mimic the reply of *P. macdermotti* females thus attracting *P. macdermotti* males, which they eat (Fig. 1b, c). When the predators are decoying *P. macdermotti* males, they sometimes produce 'extra' flashes that mimic the 'competition' flashes used by the males when competing for females (compare Fig. 1d with e)³. I suggest here that the male competition flashes (that the predators mimic) might also be mimics of the predator's own mimicry.

This complex interaction, the apparent product of a long co-evolutionary 'arms' race⁴, is more comprehensible when a model for its evolution is considered (Fig. 3). The predators also

mimic the responses of single-flash species⁵ and it is of special importance to the model that predators became tuned to *P. macdermotti*'s double flash only after they mimicked and preyed on single-flash species. Mimicry mistakes during this evolutionary transition initiated the path to the 'meta-mimicry' proposed here. The success of the tactics of each deceiver, competing male and predaceous *Photuris* female, depends on the influence the other has had, and continues to have on the system. That is, each mimicry provides the basis of support (model) for the other—such reciprocal mimicry is a new phenomenon to science.

The competition flashes (injections) that male *P. macdermotti* inject into the coded dialogues of rivals and contested females are made during the last 0.6 of the rival's flash pattern (Fig. 2b).

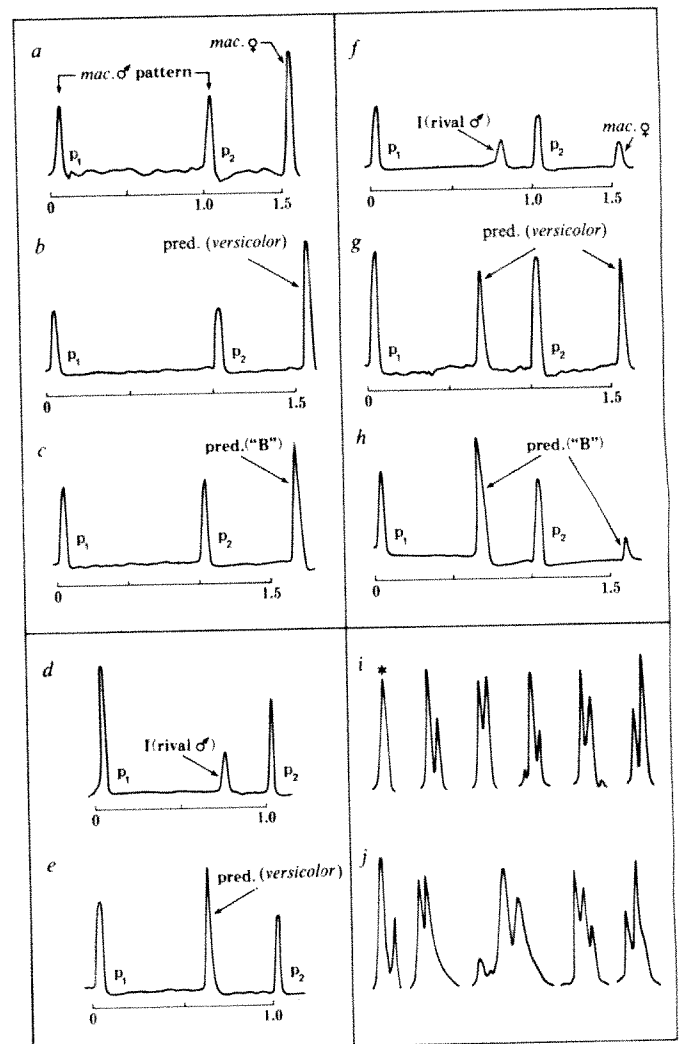


Fig. 1 Chart traces of firefly flashes and interactions, recorded with a photomultiplier system during natural and experimentally manipulated interactions in the field and laboratory. Abscissa represents time, measured as a ratio (in the text) with 1.0 being *P. macdermotti* male p_1 - p_2 interval (this removes temperature effects); ordinate, relative intensity. a, *P. macdermotti* male's characteristic two-pulse (that is, p_1 + p_2) flash pattern, and female answer. b, Simulated male pattern with aggressive mimic *P. versicolor* giving false answer. c, Same as b, with *Photuris* 'B' female giving false answer. d, Male flash pattern with rival injection (I); no female response. e, Simulated male pattern with predator *P. versicolor* giving false injection but no false answer. f, Simulated *P. macdermotti* male pattern, with injection (I) by rival male *P. macdermotti* and answer by female. g, Simulated male pattern with aggressive mimic *P. versicolor* emitting both false injection and false answer. h, Same as g, except aggressive mimic *Photuris* 'B' female. i, Various injection forms of competing *P. macdermotti* males. Note similarity to false answers (j) that aggressive mimics give to single, short flashes (see text).

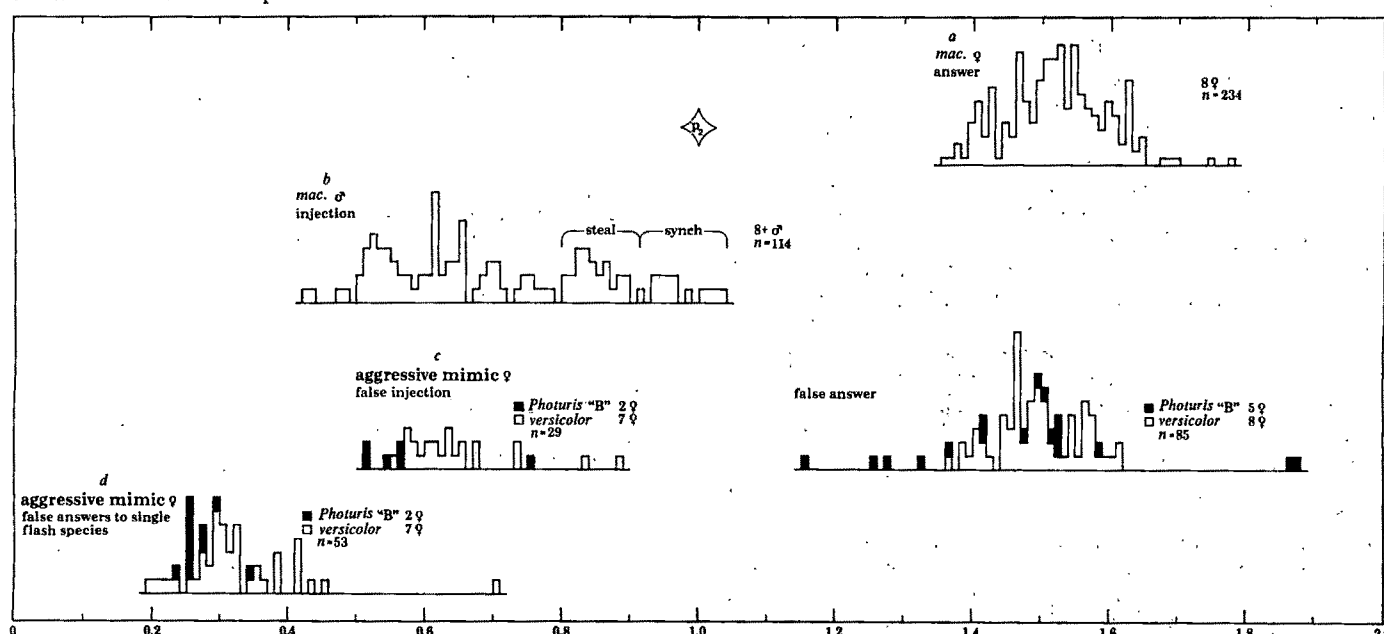


Fig. 2 Time-positioning of various flashed signals and mimics involved in the reciprocal mimicry phenomenon. Abscissa represents time, measured as a fraction of the interval between the two flashes (p_1 and p_2 in Fig. 1) of the *P. macdermotti* male pattern (this corrects for temperature differences among recording sessions as signalling parameters are temperature dependent). The first male flash (p_1) occurs at time 0; the second flash of the male pattern (p_2) occurs at time 1. *a*, Timing of *P. macdermotti* female flash responses, centering ~0.5 units after 1 (p_2). *b*, Timing of *P. macdermotti* male competition flashes (injections). 'Steal' and 'synch' indicate other male tactics³. *c*, Timing of predators' false injections (mimicking *P. macdermotti* male injections, see *b*) and false answers (mimicking *macdermotti* female answers, see *a*). Note the time correspondence between the false injections in *c* and their model in *b* above; and between false answers in *c*, and their model in *a* above. Note particularly that injections and false injections centre ~0.6 units after the preceding male flash (p_1 at time 0), but answers and false answers centre ~0.5 units after the preceding male flash (p_2 at time 1). *d*, Timing of the flashed responses used by predators when answering isolated single (non-paired) flashes—presumably the response given to some as yet undetermined, single-flash prey firefly. Note that this response centres ~0.3 time units after the 'male' stimulus flash that precedes it.

(Decimal fractions are used to indicate time-positioning because the 'clock-time' of these signals varies with temperature. Unity (1.0) is the mean period between the two flashes in the time-coded pattern of flying, searching males².) Injected flashes take various forms. Those discussed here are simple, like those in the normal flash pattern (Fig. 1*i*, asterisk), and occur between 0.5 and 0.8. The timing of these injections is similar to that of the false injections emitted by both *Photuris* predators (Fig. 2*c*). (Injections occurring later than ~0.8 seem to be involved in other competitive, photic tactics; Fig. 2*b*, 'steal' and 'synch'.)

The false injections (male imitation) and the false answers (female imitation) of predaceous females are two different responses, not merely an identical 'pause then flash' response given after different flashes of the two-flash male pattern. That is, the predators do not simply adopt the same response, but use two different pauses (that is, each matching its own model). Note that false injections (Fig. 2*c*) are centred at ~0.6 time units after the male flash that precedes them (p_1 is at 0), and false answers centre at 0.5 time units after the male flash that precedes them (p_2 is at 1.0). Note also the difference in timing of the two in the recordings (Fig. 1*g*, *h*).

The stimulus that elicits the two responses in the predators is the two-flash pattern (the species-specific pattern of *P. macdermotti* males). In contrast, if single flashes are presented at long intervals (>5 s), and therefore a *P. macdermotti*-prey sign-stimulus is not seen by the predator, a different response (pause time) is evoked—this is probably used to prey on another species. In both *P. versicolor* and 'B' this consists of a very short pause, and a modulated flash that is longer than those given in response to *P. macdermotti* patterns (Figs 1*j*, 2*d*). This demonstrates that the false injection is not merely a misplaced false-answer that is correctly used when preying on males of some single-flash species, and it further supports the conclusion that the false injection is a specifically tuned response.

The success of mimics on both sides (predator and prey) depends on the influence the other has on the system. The false injections of *Photuris* must enhance their attraction of prey

because *P. macdermotti* males use their injections to gain advantage over rivals (J.E.L., in preparation), and selection has favoured males that are not duped, but instead continue approaching with caution. The injections of competing males

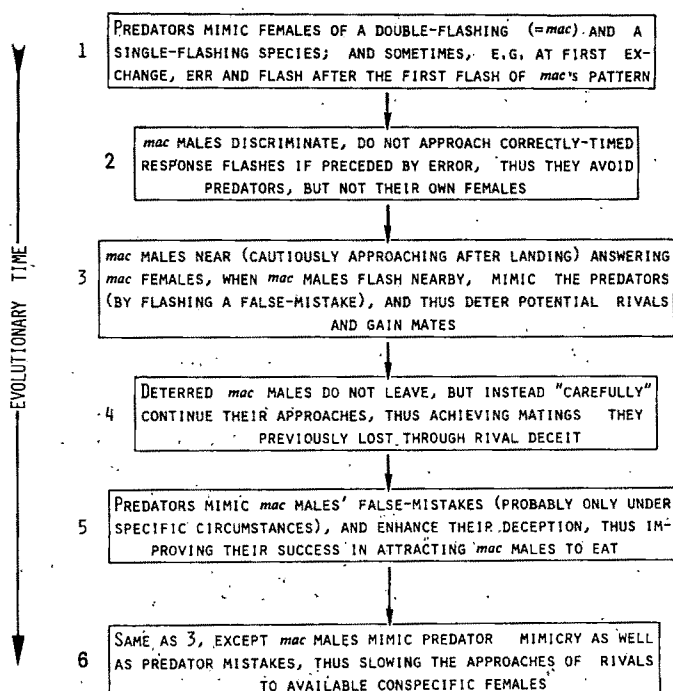


Fig. 3 Model of an evolutionary 'arms' race (conflict coevolution) leading to reciprocal mimicry of *P. macdermotti* males and their predators, *Photuris* females. A number of variants and origins are possible.

continue to influence rival behaviour because *Photuris* females continue to use them to prey on males. An individual of either species that abandons its mimicry is at a disadvantage in its own individual reproductive effort, thus natural selection independently maintains both components of this reciprocal mimicry.

Historical interactions that may have led to this dual mimicry (Fig. 3) will be discussed in more detail elsewhere, but a further possible mimicry is worth noting because it relates to the presumptive evolutionary origin of false injections, mistakes sometimes made by predators. About 75% of the injections made by *P. macdermotti* males during competition are double or triple flashes (Fig. 1i), which resemble flashes made by *Photuris* in response to single-flash stimuli (Fig. 1j). Males emit these modulated flashes only during mate competition. Their strong resemblance to flashes made by *Photuris* further suggests a highly refined mimicry system.

The communicative-deceptive transactions of these fireflies, both predators and mate-rivals, are more complex and finely tuned than expected. Similar refinement and subtlety probably exist in the communicative ecology of most insects, particularly in the contexts of female mate assessment and choice, but escape detection and general recognition because of the technical difficulties that their signals, especially chemical ones, present to discovery and analysis. Fireflies, therefore, offer unique opportunities for investigation of insect communication.

I thank J. Alcock, T. Forrest, J. Sivinski, R. Thornhill, T. J. Walker and R. Dawkins for helpful comments on the manuscript and S. Wineriter for preparing the illustrations. This work was supported by Departmental and personal funds. It is Florida Agricultural Experimental Station Journal Series no. 2304.

Received 17 July 1980; accepted 13 January 1981.

1. Lloyd, J. E. A. *Rev. Ent.* **16**, 97–122 (1971); in *Bioluminescence in Action* (ed. Herring, P. J.) 241–272 (Academic, New York, 1978).
2. Lloyd, J. E. *Univ. Mich. Mus. Zool. Misc. Publ.* **130**, 1–95 (1966); *Coleopt. Bull.* **23**, 29–40 (1969).
3. Lloyd, J. E. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. S. & Blum, N. A.) 293–342 (Academic, New York, 1979).
4. Dawkins, R. & Krebs, J. R. *Proc. R. Soc. B205*, 485–511 (1979).
5. Lloyd, J. E. *Science* **149**, 653–654 (1965); **187**, 452–453 (1975); *Coleopt. Bull.* **27**, 91–106 (1973).

Removal of bound calcium from nematocyst contents causes discharge

Roger Lubbock & W. B. Amos

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

Nematocysts are the stinging organelles of jellyfish, sea anemones and other cnidarians. Each one consists of a closed capsule filled with fluid. In the resting state, part of the surface of the capsule is inverted, forming a tubular thread, which is everted explosively on excitation. The mechanism of explosion is not yet understood^{1–3}; it may be relevant to exocytosis in general, as the nematocyst is a specialized type of exocytotic secretion originating in the Golgi apparatus⁴. Picken and Skaer¹ observed that the capsular fluid showed a large depression of freezing point, suggesting that the osmotic pressure might be as high as 140 bar. They tentatively ascribed the explosion to a sudden increase in permeability of the capsule wall, allowing a rapid osmotic influx of water. However, there is evidence that the material of the thread itself may be capable of some degree of extension^{5,6} and it has been suggested that osmosis plays no part in discharge³. We present here a new theory: that the capsule wall is permeable to water even in the undischarged state; discharge is initiated by an increase in the osmotic pressure of the capsular fluid which is brought about by removal of bound calcium ions.

The work was carried out on holotrichous isorhizas from the mesenterial filaments of the coralliomorpharian anemone *Rhodactis rhodostoma*. These nematocysts have the advantage that they are large enough (150 μm long) to be manipulated individually with fine steel needles. All manipulations were carried out in liquid paraffin. Before use, the nematocysts were isolated from the mesenterial filaments by teasing with needles and cleaned of mucus and cell debris by drawing them carefully across a glass surface. Experimental solutions were injected as droplets into the liquid paraffin and the nematocysts transferred from one droplet to the next using a steel needle. Measurements of refractive index were carried out with a Zeiss Jamin–Lebedeff system and with a Baker image-shearing interference system equipped with a fringe-field eyepiece. The freezing point depression of the contents of intact individual capsules was measured by a method to be described in detail elsewhere. It consisted essentially of placing the isolated nematocysts in liquid paraffin inside a sealed chamber which was immersed in a refrigerant of a Ramsay–Brown freezing point apparatus¹. The chamber was viewed microscopically and the formation and melting of ice crystals within individual capsules observed.

This new method of *in situ* measurement revealed slightly higher freezing point depressions than those observed by Picken and Skaer¹, who used nematocysts of *Corynactis* and determined the freezing point of extracted samples of fluid. Our *in situ* measurements on *Rhodactis* gave values for $\Delta^\circ\text{C}$ of -5.4 to -8.2 ($\bar{x} = 6.2$, $n = 15$), whereas those of Picken and Skaer were -4.2 to -5.9 . These values correspond to high notional osmotic pressures, but, as mentioned by Wolf⁷, there is no simple relationship between osmotic pressure and freezing point depression for concentrated macromolecular solutions. Without further knowledge of the chemical composition of the capsular fluid, the osmotic pressure cannot be deduced.

As previously observed⁸, the capsule fluid had a high refractive index. Undischarged capsules gave values as high as 1.378–1.388, corresponding to dry mass concentrations of 25–30% ($n = 4$) in excess over seawater, assuming a refractive increment of 0.0018 (ref. 9). The latter assumption is valid if the capsule solutes are protein, glycoprotein, carbohydrate or amino acids. Capsules which discharged spontaneously underwent a dramatic drop in refractive index over several seconds, which caused the specimens to pass through four orders of interference colour. The final value of refractive index corresponded to a mean dry mass excess over seawater of only 0.56% (s.d. ± 0.16 , $n = 10$) which is probably due to the mass of the capsule wall. It is therefore clear that during discharge there is a massive influx of seawater into the capsule, as observed by others, and that the capsule contents are completely replaced. These observations show that a concentrated and probably highly osmotically active solution in the capsule is drastically diluted during discharge; thus far our findings are consistent with the theory of Picken and Skaer.

Further experiments were performed in an attempt to discover the nature of the process that initiates the influx of water. Attention was directed to calcium because microprobe measurements have shown that the capsule fluid contains calcium at a concentration of about 600 mM per kg wet weight¹⁰. Isolated nematocysts could be made to discharge by placing them in a droplet containing 10 mM potassium citrate¹¹ with 10 mM imidazole buffer, pH 7.0. A similar solution without citrate was ineffective. The excitatory effect was blocked by the addition of 50 mM CaCl_2 to the citrate/imidazole solution (only 1 capsule out of 15 discharged). This blockage was peculiar to calcium, other cations being ineffective: 15 capsules out of 15 discharged in each case when NaCl , KCl or MgCl_2 was used as 50 mM solutions in place of CaCl_2 . The stimulatory effect of citrate is almost certainly due to its chelation of calcium; similar effects were obtained with EGTA. This experiment indicates that calcium ions stabilize the nematocyst in some way (as suggested by Blanquet¹¹) and moreover that their removal may be the cause of discharge.

Table 1 The effects of CaCl_2 and sodium citrate on the tonicity of undischarged capsules

Experiment	No. of shrunken capsules	
	1st solution	2nd solution
CaCl_2 followed by sodium citrate	5	0
CaCl_2 followed by CaCl_2	5	5
Sodium citrate followed by CaCl_2	0	5
Sodium citrate followed by sodium citrate	0	0

In each experiment, groups of five undischarged capsules were placed sequentially for 1 min each in two separate solutions containing 40% w/v polyethylene glycol, 50 mM imidazole buffer, pH 7.0, and either 50 mM CaCl_2 or 50 mM sodium citrate; the state of expansion of the capsules was recorded. A repeat of these experiments with 50 mM MgCl_2 added to the citrate solutions gave similar results. Previously discharged capsules containing only seawater shrank in all the above solutions.

One possibility was that calcium might prevent discharge by rendering the capsule wall impermeable to water. To test this hypothesis, undischarged capsules were placed in a series of stain solutions containing 50 mM imidazole buffer, pH 7.0, and 50 mM CaCl_2 . Capsules were found to be impermeable to a 5% solution of blue dextran (molecular weight (MW) 2×10^6) or to 1% Alcian blue (MW 1,265), even when left for several minutes. However, 1% solutions of phenol red (MW 354), sodium fluoresceinate (MW 376) and bromophenol blue (MW 670) were all able to permeate the capsule, reaching approximately the same concentration as outside within 10–20 s; capsules that had been permeated with stain retained their ability to discharge when placed in a buffered 10 mM citrate solution. These results show that the wall of the undischarged capsule is permeable to charged molecules of relatively low molecular weight. It therefore seems likely that very small molecules such as water can diffuse rapidly in and out of the undischarged nematocyst.

We have considered two other possible mechanisms for the action of calcium. One is that the osmotic pressure of the internal fluid is constant and is continuously exerted; discharge is then made possible by a localized weakening of part of the thread or capsule wall, due perhaps to the loss of calcium as a cross-linking agent. This seems unlikely, as the capsule expands briefly before discharge, as previously observed³, indicating an increase in hydrostatic pressure. An alternative mechanism is that the removal of calcium changes the nature of the capsule fluid such that the osmotic pressure increases.

The latter hypothesis was tested by measuring the tonicity of undischarged capsules with and without calcium in the bathing medium. To prevent discharge, it was necessary to use a solute that could not readily penetrate the capsule wall but was of low enough molecular weight to allow an external solution of relatively high osmotic pressure; polyethylene glycol of MW 1,500 was chosen. Undischarged nematocysts were placed in solutions containing 40% w/v polyethylene glycol, 50 mM imidazole buffer, pH 7.0, and either 50 mM CaCl_2 or 50 mM sodium citrate (Table 1). The capsules shrank when placed in the calcium solution, but re-expanded when placed in the citrate solution. In the reverse experiment, capsules placed first in citrate remained expanded, but shrank when transferred to the calcium solution. These results cannot be ascribed to the increase in the osmotic pressure of the bathing medium due to the addition of the calcium salt, as the addition of 50 mM MgCl_2 to the citrate solutions failed to cause shrinkage. Therefore, these experiments strongly suggest that the tonicity of the capsule fluid in the presence of calcium ions is lower than in their absence.

We propose that the isolated capsule is continuously permeable to water and the internal hydrostatic pressure due to osmosis is sufficient to keep it turgid but insufficient to produce discharge. The normal stimulus for discharge is the removal of calcium from the nematocyst fluid, presumably in ionic form. This results in an increase in the osmotic pressure of the fluid.

The molecular basis of this process is not yet clear; it could involve the dissociation of a macromolecule, creating a larger number of osmotically active particles.

Received 10 December 1980; accepted 4 February 1981.

1. Picken, L. E. R. & Skaer, R. J. *Symp. zool. Soc. Lond.* **16**, 19–50 (1966).
2. Robson, E. A. *Publ. Seto mar. biol. Lab.* **20**, 653–673 (1973).
3. Mariscal, R. N. in *Coelenterate Biology: Reviews and New Perspectives* (eds Muscatine, L. & Lenhoff, H.) 129–178 (Academic, London, 1974).
4. Skaer, R. J. *J. Cell Sci.* **13**, 371–393 (1973).
5. Robson, E. A. *Q. Jl microsc. Sci.* **94**, 229 (1953).
6. Carré, D. *Eur. J. Cell Biol.* **20**, 265–271 (1980).
7. Wolf, A. V. *Aqueous Solutions and Body Fluids: Their Concentrative Properties and Conversion Tables* (Harper & Row, New York, 1966).
8. Sernetz, M. & Schmidt, H. *Microsc. Acta* **73**, 109–116 (1973).
9. Ross, K. F. A. *Phase Contrast and Interference Microscopy for Cell Biologists* (Edward Arnold, London, 1967).
10. Lubbock, R., Gupta, B. L. & Hall, T. A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
11. Blanquet, R. *Comp. Biochem. Physiol.* **35**, 451–461 (1970).

Immunoradiometric assay of human leukocyte interferon using monoclonal antibody

D. S. Secher

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Interferon research has been hampered by problems in its assay. The only widespread assays use tissue culture cells and compare some parameter of viral growth (such as viral RNA synthesis or host cell death) in the presence and absence of interferon¹. These complex biological assays, though sensitive, are laborious and subject to inherent variability. In particular, components other than interferon present in the assay sample often influence viral growth. There is clearly a need for a simple, direct, interferon assay; I now describe such an assay—an immunoradiometric assay utilizing monoclonal antibody.

The isolation and properties of a monoclonal antibody (NK2) specific for human leukocyte interferon (Hu-IFN α) were recently reported². This antibody was used to purify Hu-IFN α by immuno-adsorption chromatography. The use of labelled antibody (immunoradiometric assay) rather than labelled antigen (radioimmunoassay) may offer improved assay sensitivity for the reasons given by Miles and Hales³ for insulin assay. Monoclonal antibodies are much more easily purified than are monospecific antibodies from conventional antisera and so will allow the widespread exploitation of the advantages of immunoradiometric assays. For this reason, it was of interest to develop a labelled antibody sandwich assay, in which a sheep anti-interferon antibody⁴ is attached to polystyrene and serves to anchor the interferon present in the sample to the solid phase. The bound interferon is then detected by the addition of ¹²⁵I-labelled NK2 (monoclonal anti-Hu-IFN α) and measurement of the counts bound to the solid phase.

Three forms of polystyrene were used as the solid phase: (1) 3-ml test tubes (LP3, Luckham); (2) 96-well microtitre trays (M24, Gibco); and (3) 6.5-mm beads (Northumbria Biologicals). In the first case the whole tube was counted to measure the bound ¹²⁵I-NK2. When trays were used the bottom of each well was cut off with a hot wire and transferred to a clean tube for counting. In the third case the beads were incubated in 20- or 60-well trays (93-0402, 93-0340, Abbott) and transferred to tubes for counting. The incubation volumes used were 1 ml or 0.1 ml (tubes), 0.1 ml (wells) and 0.2 ml (beads). All three supports were satisfactory and for assaying large numbers of samples the beads were preferred for their convenience.

NK2 antibody was purified from the serum and ascites fluid of mice carrying NK2 tumours by ammonium sulphate precipitation and DEAE-ion-exchange chromatography, labelled by the chloramine-T method and desalted on a Sephadex G-50

(fine) column. The labelled IgG had a specific activity of about $2 \text{ Ci } \mu\text{mol}^{-1}$ or about 1 atom ^{125}I per molecule IgG.

In each assay a standard curve was constructed using either the interferon reference standard MRC 69/19 or a laboratory standard. Such a curve is shown in Fig. 1. The counts bound at $8,000 \text{ U ml}^{-1}$ were at a maximum; no further increase was observed with interferon concentrations up to 10^6 U ml^{-1} . The low level of nonspecific binding ($<1\%$) of input counts is an important characteristic of the sandwich assay.

The assay conditions (Fig. 1 legend) were chosen for the convenient measurement of samples encountered in the purification of interferon, with concentrations of $\sim 10^3$ – 10^7 U ml^{-1} . Solutions containing $>10^3 \text{ U ml}^{-1}$ are diluted by serial dilution and the titration curve obtained matched to the standard curve.

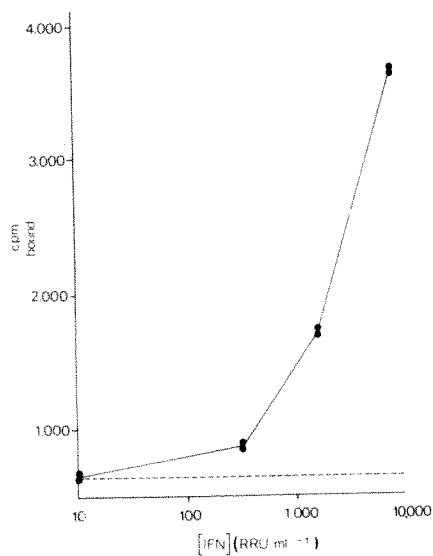


Fig. 1 Standard curve for measurement of interferon concentration by immunoradiometric assay (as obtained by J. Walker). IgG was purified from a sheep anti-Hu-IFN α antiserum⁴ (450,000 neutralizing units per ml) by ammonium sulphate precipitation and DEAE-cellulose ion-exchange chromatography and coated on polystyrene beads by incubation of the beads at 4°C for 16 h in the sheep IgG ($20 \mu\text{g ml}^{-1}$ in PBS, 5 mM EDTA, 0.1% NaN_3). Several hundred beads were coated, washed in PBS, 10% horse serum (Sera-Lab), 0.1% NaN_3 (HS medium) and stored in HS medium at 4°C . Assay trays (20- or 60-well, Abbott) were incubated with HS medium at 4°C to block any sites for protein attachment. Interferon samples were diluted in HS medium and duplicate samples ($200 \mu\text{l}$) added to antibody-coated beads (1 bead per well). After 4 h at 4°C the beads were washed with 12 ml HS medium each using a combined dispenser/aspirator (Pentawash, Abbott). After removal of any residual medium, ^{125}I -NK2 (purified IgG, 40,000 c.p.m.) was added and incubated for 16 h at 4°C . The beads were washed as before and transferred to a γ -counter. The dotted line indicates the background c.p.m. bound when the coated beads were incubated in HS medium without interferon. The IFN concentration is expressed in Reference Research Units (RRU) per ml using the MRC standard 69/19 (National Institute for Biological Standards and Control).

The immunoradiometric assay offers considerable advantages over the conventional biological assays. Since the antibody-coated polystyrene can be stored at 4°C , samples can be assayed at short notice and the results are obtained within 24 h of beginning the assay. The reproducibility, between assays, of the immunoradiometric assay is much better than that of biological assays; four measurements of a solution of $2,000 \text{ U ml}^{-1}$ gave a value of $3,063 \pm 345$ c.p.m. in independent assays in which the input c.p.m. was 47,000–53,000 c.p.m. Only small amounts of sheep antibody are used, since the IgG solution used to coat the polystyrene can be used several times without apparently reducing the sensitivity of the assay. The quality of the sheep antibody may not be a critical factor in the success of the assay,

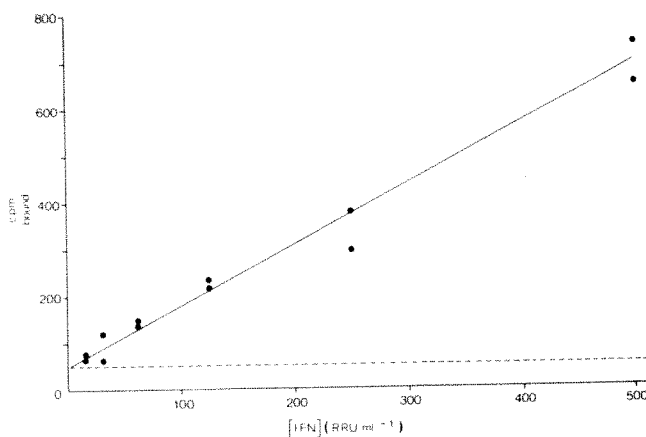


Fig. 2 Standard curve for low levels of interferon. Interferon samples were prepared by dilution of the reference research standard in HS medium and assayed as in Fig. 1 legend except that only 23,000 c.p.m. ^{125}I -NK2 was added to each bead. The dotted line indicates the background c.p.m. bound and was the mean of eight values (s.e.m. = ± 6 c.p.m.).

and other antibodies to Hu-IFN α could probably be substituted. The assay exploits the specificity of the monoclonal antibody NK2, and because this is the product of a hybrid myeloma cell line in culture it can be produced in large amounts without any change in quality. Finally the assay is inexpensive (especially since all the plastic ware except the beads may be re-used), lends itself to automation and several hundred samples can be assayed by one person in a day. Antiviral assays will still be needed to detect any interferons that do not possess the NK2 antigenic site (including IFN β (ref. 5) and possibly some members of the IFN α family) and to confirm biological activity in antigenically positive molecules.

For the measurement of interferon in biological fluids an assay must be capable of detecting much lower levels of interferon than those encountered in production and purification. Figure 2 shows that at low interferon concentrations the counts bound are proportional to the interferon concentration, and concentrations of $\geq 50 \text{ U ml}^{-1}$ can be reliably measured.

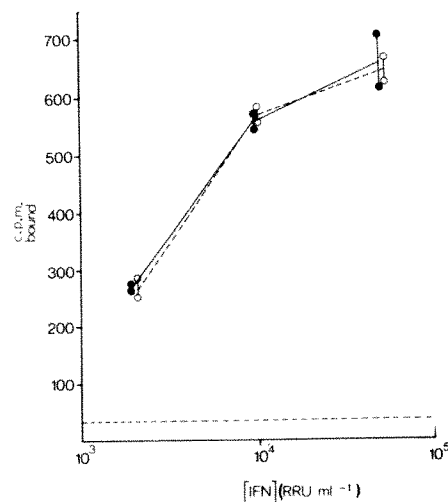


Fig. 3 Interferon assay in the presence of human serum. LP3 tubes were coated with sheep anti-interferon ($100 \mu\text{l}$ per tube, conditions as in Fig. 1 legend). Lymphoblastoid interferon (from stock of 10^5 U ml^{-1}) was diluted in HS medium (●) or undiluted normal human serum (○) and $100 \mu\text{l}$ samples assayed in duplicate by incubating at 4°C for 4 h, washing with $3 \times 2 \text{ ml}$ HS medium and then incubating for a further 16 h at 4°C in the presence of 35,000 c.p.m. ^{125}I -NK2 IgG ($100 \mu\text{l}$). After washing again the tubes were counted for radioactivity.

A limitation in the conventional assay of interferon in serum or other biological fluids has been the need to dilute the sample until the effects of other non-interferon substances that affect viral growth no longer mask the action of the interferon. Dilution of a sample of lymphoblastoid interferon in either phosphate-buffered saline (PBS), 10% horse serum, 0.1% NaN_3 or in undiluted normal human serum resulted in identical titration curves (Fig. 3), indicating that the immunoradiometric assay overcomes this problem.

The lower limit in the measurement of serum interferon by biological assays has been sufficient to detect interferon in the serum of some patients with acute viral infections and in the serum of patients to whom interferon has been administered. This sensitivity is insufficient, however, to measure the interferon (if any) in normal human serum.

It should be possible to improve the sensitivity of the immunoradiometric assay still further by optimizing the conditions at low interferon concentrations. In any case the sensitivity of the assay can be further increased by prior concentration and purification of the interferon on a small NK2-affinity column.

I thank Dr K. Cantell for sheep anti-interferon antiserum, Dr Cantell and Professor D. C. Burke for gifts of interferon and for critical discussions, Drs J. Nagington and J. Walker for collaborative studies to be published separately and for introducing me to the coated bead system of radioimmunoassay, Dr C. Milstein for suggestions and encouragement, and R. E. Hawkins and L. T. Davies for technical assistance.

Received 12 September 1980; accepted 19 February 1981.

1. Finter, N. B. in *Interferons and Interferon Inducers* (ed. Finter, N.B.) 135-170 (North-Holland, Amsterdam, 1973).
2. Secher, D. S. & Burke, D. C. *Nature* **285**, 446-450 (1980).
3. Miles, L. E. M. & Hales, C. N. *Nature* **219**, 186-189 (1968).
4. Mogensen, K. E., Pyhälä, L. & Cantell, K. *Acta path. microbiol. scand.* **B83**, 443-450 (1975).
5. Morser, J., Meager, A., Burke, D. C. & Secher, D. S. *J. gen. Virol.* (in the press).

Evidence that the serological determinant of H-Y antigen is carbohydrate

Mark Shapiro & Robert P. Erickson

Department of Human Genetics, Box 015, University of Michigan Medical School, 1137 E. Catherine Street, Ann Arbor, Michigan 48109, USA

The histocompatibility-Y (H-Y) antigen is a minor histocompatibility antigen which has been detected on cell surfaces from the heterogametic sexes of mammalian, bird, amphibian, teleost and invertebrate species¹⁻³. H-Y is thought to be a male-determining substance in mammals because of its almost perfect correlation with maleness among a variety of mammalian species. To characterize the molecular determinant responsible for H-Y-specific serological activity, H-Y-positive immunoabsorbent cells were first subjected to various treatments which alter protein or carbohydrate structure and then tested for their ability to absorb H-Y antisera. We present here evidence that the serological determinant of H-Y antigen is carbohydrate.

To assess the dependence of H-Y serological activity on polypeptide structure, absorbing cells were treated with the proteases trypsin, α -chymotrypsin and thermolysin. After enzyme treatment and washing, treated splenocytes were used to absorb an aliquot of H-Y antiserum. In this and all subsequent experiments, absorbed sera were tested for cytotoxic activity against male murine epidermal cells. Splenocytes treated with either trypsin, α -chymotrypsin or thermolysin

showed almost the same absorptive capacities as untreated splenocytes (Fig. 1). These results indicate that H-Y-specific serological activity is resistant to these three proteolytic enzyme treatments.

To determine the dependence of H-Y serological activity on carbohydrates, absorbing cells were treated with periodate, which oxidizes 1,2-dihydroxyethyl linkages found in glycosyl residues. Splenocytes were treated with varying concentrations of periodate and washed in a protein+sugar-supplemented medium before absorption. Sensitivity of H-Y-specific absorption to periodate treatment, as well as subsequent experimental treatments, were quantified as corrected per cent inhibition of absorption, as defined in Fig. 2 legend. As a control, the same treatments were performed on absorbing cells bearing antigens of the H-2^b haplotype. Such serological determinants are thought to be contained in amino acid sequences^{4,5}. These data show H-Y-specific absorption to be ~ fivefold more periodate sensitive than H-2 (Fig. 2a).

A more direct assessment of the carbohydrate dependence of H-Y serological activity can be made by treating absorbing cells

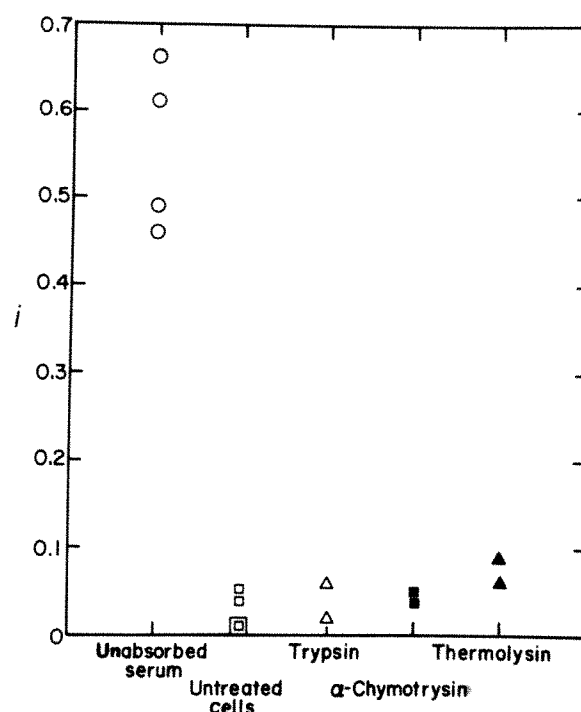


Fig. 1 H-Y-specific absorption: resistance to proteolytic enzymes. Purified trypsin (Gibco), α -chymotrypsin (Worthington), and thermolysin (Calbiochem) were dissolved in Puck's saline F (PSF) to yield individual suspensions containing 0.1% enzyme by weight. Samples containing 4×10^7 splenocytes were incubated in a 2-ml enzyme suspension at pH 7.4, 37°C for 20 min, washed three times at 4°C in Medium 199 containing 10% heat-inactivated (56°C, 30 min) γ -globulin-free fetal calf serum (FCS, Gibco) and used to absorb 25- μ l aliquots of rat H-Y antiserum diluted 1:4 in Medium 199 containing 10% FCS. Rat H-Y antiserum was prepared by the method of Fellous *et al.*⁹ by immunizing female Fisher rats with male (Fisher) splenocytes. Before use in this experiment, rat antisera were heat inactivated (56°C, 30 min) and absorbed with equal volumes of female murine splenocytes for 1 h at both 0 and 37°C. An antiserum was regarded as H-Y specific if it had cytotoxic activity against male, but not female, C57BL/6J (B6) epidermal cells, and this activity could be absorbed by male, but not female B6 splenocytes. H-Y-specific immunoabsorption was done on ice (~0°C) for 1 h. Absorbed sera were tested for direct cytotoxicity against B6 epidermal cells according to the method of Scheid *et al.*¹⁰ with the exception that heat-inactivated γ -globulin-free FCS was used as a protein supplement for the isolation medium. Cell killing was quantified by the cytotoxic index, *i*, defined as: no. dead cells in sample - no. cells killed by complement dose/total no. cells - no. cells killed by complement alone.

with exo-glycosylhydrolases which remove a terminal glycosyl residue from the non-reducing end of a carbohydrate chain. Splenocytes were treated with highly purified β -glucuronidase and β -galactosidase, then washed in protein-enriched media before absorption. Figure 2b shows a graph of corrected per cent inhibition of absorption versus units of enzyme used for treatment. Whereas H-Y-specific absorption is resistant to β -glucuronidase treatment (1,000 units of enzyme destroy 2% of activity from the absorbing cells), it is sensitive to β -galactosidase to the extent that 1,000 units of enzyme remove 54% of the H-Y absorptive capacity of male mouse splenocytes. To confirm that the inhibition by β -galactosidase is not due to uptake of the enzyme by the absorbing cells, the 1,000-unit enzyme treatment was repeated for H-Y in the presence of 0.1 M mannose-6-phosphate, which inhibits the uptake of β -galactosidase by mammalian cells. In the presence of mannose-6-phosphate, the 1,000-unit β -galactosidase treatment still destroyed 45% of H-Y-specific absorptive activity (Table 1). As another control, the 1,000-unit β -galactosidase treatment was used on absorbing cells specific for H-2^b antigens. H-2^b-specific

absorption is completely insensitive to the action of 1,000 units of β -galactosidase (Table 1).

These data suggest that the integrity of an H-Y serological determinant is perturbed by the removal of a terminal galactose from the non-reducing end of an associated carbohydrate moiety. To test further the serological dependence of H-Y on a non-reducing terminal galactose residue, we performed a series of absorbing cell treatments using galactose oxidase, which oxidizes the terminal galactose residue at the non-reducing ends of a carbohydrate chain. Although galactose oxidase has no appreciable effect on H-2^b-specific absorption, it inhibits H-Y-specific absorption by male mouse splenocytes by as much as 38% in this experiment (Fig. 2c).

To verify that the inhibition of H-Y observed here is due to the specific oxidation of a terminal non-reducing galactose rather than to nonspecific oxidation at the cell surface, a similar experiment was carried out using glucose oxidase. Whereas 10 units of galactose oxidase inhibit H-Y-specific absorption by 38%, an equivalent amount of glucose oxidase had only an 8% inhibitory effect (Table 1). This value could possibly be due to

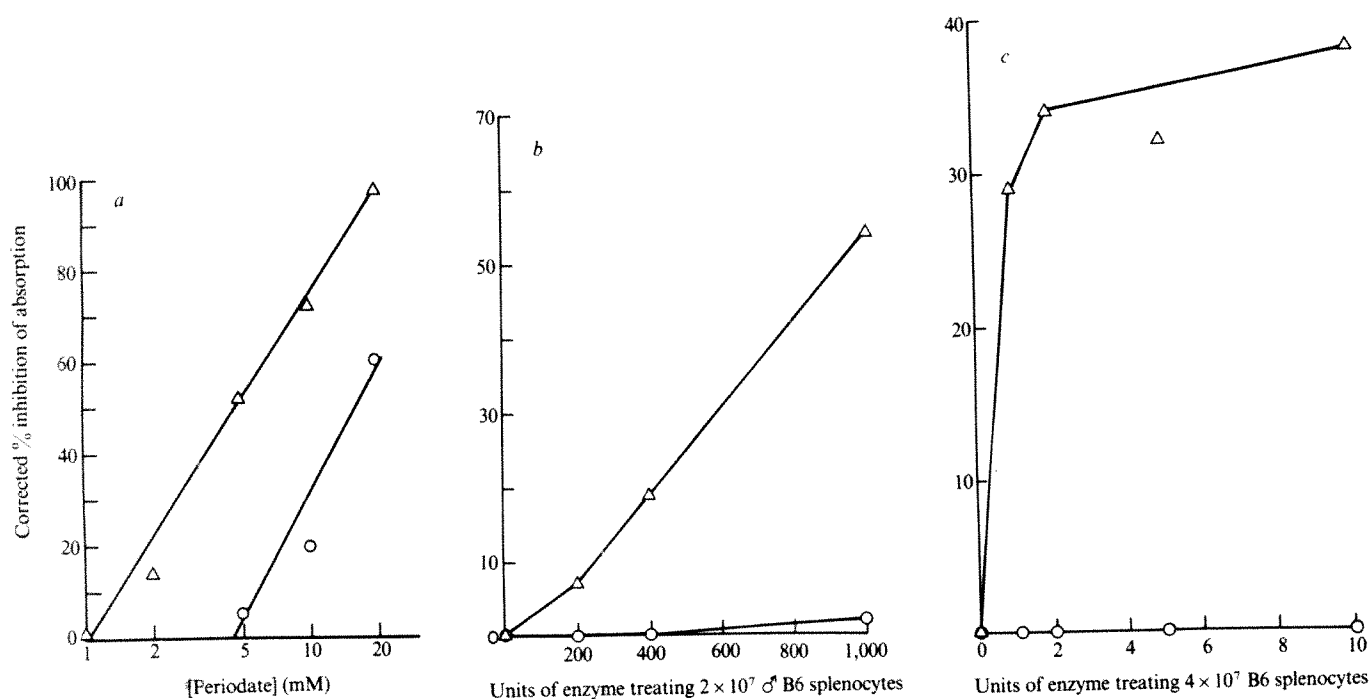


Fig. 2 Sensitivity of H-Y-specific absorption to all treatments quantified as corrected per cent inhibition of absorption, defined as:

$$\frac{i(\text{treated cells}) - i(\text{untreated cells})}{i(\text{unabsorbed serum}) - i(\text{untreated cells})}$$

with i defined as in Fig. 1 legend. Each per cent inhibition value represents the mean of duplicate samples. **a**, Known concentrations of periodic acid (Frederick Smith Chemicals) were dissolved in isotonic phosphate buffer and adjusted to pH 7.4. Splenocytes were treated at a concentration of 10^7 cells per ml (10^7 total cells for H-2 and 2×10^7 total cells for H-Y) for 10 min at room temperature ($\sim 22^\circ\text{C}$). After treatment, the cells were washed twice in Medium 199 containing 10% FCS at 4°C . Treated splenocytes were used to absorb either mouse H-Y (diluted 1:4) or H-2^b antiserum (diluted 1:100) as described in Fig. 1 legend. Mouse H-Y antiserum was prepared as described by Krco and Goldberg⁴. H-2^b antiserum was an equivolume mixture of alloantisera to murine H-2.2(D-2) and H-2.33(D-33) specificities obtained from the NIH. Absorbed H-Y antiserum was tested for cytotoxicity as described above. H-2^b antisera were tested for direct cytotoxicity against splenocytes in the following way. Equal amounts (10 μl) of H-2^b antiserum and B6 splenocytes (8×10^6 per ml) were incubated for 20 min at room temperature; 10 μl of baby rabbit complement (Pel Freez) diluted 1:4 was then added and the suspension incubated at 37°C for 35 min. Trypan blue (10 μl of 0.2% dye) was added for the final 5 min of incubation to stain dead splenocytes. Formalin saline (10 μl containing 5% formaldehyde) was added as a fixative and live and dead cells were counted. $\circ$, H-2^b; Δ , H-Y. **b**, Highly purified rat β -glucuronidase ($\circ$) and bovine β -galactosidase (Δ) (both 12,000 units per mg protein) were obtained from Dr Virginia Hieber, Department of Pediatrics and Communicable Diseases, University of Michigan Medical School. Splenocytes (2×10^7 total) were treated with known amounts of enzyme in 2.0 ml incubation volumes containing phosphate buffered saline (PBS) at pH 6.8, 37°C for 1 h. After treatment, cells were washed twice in Medium 199 containing 10% FCS at 4°C before absorption. Antisera used for absorption and subsequent cytotoxicity testing were murine H-2^b and rat H-Y antisera as described above. **c**, Galactose oxidase (Sigma) was dissolved in PBS containing 1.0 mg per ml of bovine serum albumin (BSA, Sigma). Splenocytes were incubated with known amounts of enzyme in PBS, pH 7.4, at 37°C for 1 h. Each incubation mixture contained 10^7 (in the case of H-2^b, $\circ$) or 4×10^7 (in the case of H-Y, Δ) splenocytes plus enzyme at a volume of 0.5 ml. After incubation, treated cells were washed in Medium 199 supplemented with 1.0 M galactose before absorption. Antisera were either murine H-2^b or rat H-Y as described above. Absorption and subsequent cytotoxicity testing were as before.

Table 1 Specificities of enzyme treatment

Enzyme treatment	Total enzyme (U)	% Inhibition of absorption (mean $\pm$ s.e.)	
		H-2 ^b	H-Y
β -Galactosidase alone	1,000	0 $\pm$ 1	54 $\pm$ 8
β -Galactosidase + 0.1 M mannose-6-phosphate	1,000	—	45 $\pm$ 4
Galactose oxidase	10	0 $\pm$ 1	38 $\pm$ 2
Glucose oxidase	10	0 $\pm$ 1	8 $\pm$ 2

β -Galactosidase treatments in the presence of mannose-6-phosphate (Sigma) were as described for Fig. 2b with the exception that 0.1 M mannose-6-phosphate was added to the treatment buffer. Glucose oxidase treatments were as described for Fig. 2c with the exceptions that glucose oxidase (Sigma) was the enzyme used and that treated cells were washed in Medium 199 supplemented with 1.0 M glucose.

the presence of contaminating amounts of galactose oxidase in our enzyme preparation (2% according to manufacturer). In addition, H-2^b was tested for inhibition of absorption by glucose oxidase and no effect was observed (Table 1). These data are consistent with the suggestion that the integrity of an H-Y serological determinant is perturbed by structural alterations directed against a terminal non-reducing galactose on an associated carbohydrate moiety.

The serological reactivity of H-Y antigen, as measured by immunoabsorption, is demonstrated here to be resistant to the action of various proteolytic enzyme treatments. This result implies that the H-Y serological determinant is not a protein, or that it is part of a polypeptide chain resistant to proteolysis. The resistance to proteolysis of H-Y, in combination with its greater sensitivity to periodate than H-2 suggest that the H-Y-specific serological determinant is more likely to be carbohydrate than protein. In most of the experiments reported here, H-2 is used as a control as it is a well characterized glycoprotein antigen whose serological activity is resistant to treatments directed at carbohydrate structures, but sensitive to protein-directed treatments^{5,6}.

We have reported that H-Y-specific absorption is sensitive to both β -galactosidase and galactose oxidase treatments. The sensitivity of H-Y-specific absorption to both these enzymes is in the range 29–54%. Although these values do not reflect complete inhibition, they are consistent with the accepted model of antigen-antibody interaction involving binding of the combining site of an antibody molecule with a carbohydrate determinant of 4–6 glycosyl residues⁷. By perturbing a terminal galactose residue on such a carbohydrate determinant, one would expect detectable, although not complete, inhibition of antibody binding. Thus, our observations of β -galactosidase and galactose oxidase sensitivity indicate a role for a terminal non-reducing galactose residue in the serological integrity of H-Y antigen.

If the serologically active portion of H-Y is carbohydrate, this could account for its evolutionary conservation, suggested by the broad serological cross-reactivity of H-Y-like substances from a variety of species^{1,2,8}. A carbohydrate nature for this determinant could also explain why antisera raised against it are of low titre (1:16–1:32 in the case of our antisera) and affinity. Although the results reported here suggest a carbohydrate nature for the H-Y serological determinant, direct demonstration of whether this immunological determinant is itself a carbohydrate, or is strictly dependent on a carbohydrate chain requires further experimentation.

This work was presented at the meeting of the American Society for Cell Biology in Cincinnati, Ohio on 14–18 November 1980. The abstract of this presentation is published elsewhere⁸.

We thank Dr Virginia Hieber for her gift of purified β -glucuronidase and β -galactosidase and Mrs Rena Jones for secretarial assistance. This work was supported by grant HD 11738 and predoctoral training grant GM 7544 from the NIH.

Received 2 January; accepted 17 February 1981.

- Ohno, S. *Major Sex Determining Genes* (Springer, Berlin, 1979).
- Pechan, P., Wachtel, S. S. & Reinboth, R. *Differentiation* **14**, 189–192 (1979).
- Shalev, A., Goldenberg, P. Z. & Huebner, E. *Differentiation* **16**, 77–80 (1980).
- Krcso, C. J. & Goldberg, E. H. *Science* **193**, 1134–1135 (1976).
- Nathanson, S. G. & Maramatsu, T. in *Glycoproteins of Blood Cells and Plasma* (eds Jamieson, G. A. & Greenwalt, J.) 245–262 (Lippincott, Philadelphia, 1971).
- Pancake, S. J. & Nathanson, S. G. *J. Immun.* **111**, 1086–1092 (1973).
- Day, E. D. *Advanced Immunocytochemistry* 123–177 (Williams & Wilkins, Baltimore, 1972).
- Shapiro, M. & Erickson, R. P. *J. Cell Biol.* **87**, 135 (1980).
- Fellous, M. *et al. J. exp. Med.* **148**, 58–70 (1978).
- Scheid, M., Boyse, E. A., Carswell, E. A. & Old, L. J. *J. exp. Med.* **135**, 938–955 (1972).

Fibrinolysis with acyl-enzymes: a new approach to thrombolytic therapy

R. A. G. Smith, R. J. Dupe, P. D. English & J. Green

Beecham Pharmaceuticals Research Division, Biosciences Research Centre, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ, UK

Deep vein thrombosis in man presents a considerable clinical challenge. Despite the availability of prophylactic measures, therapeutic thrombolysis is often necessary, but is difficult and hazardous. Treatments have included the administration of plasmin, other less specific proteolytic enzymes, the indirect plasminogen activator, streptokinase, and the direct activators, urokinase and streptokinase-human plasmin complex. All these treatments have been associated with some haemostatic breakdown, which has discouraged their widespread application. The enzyme components of the coagulation and fibrinolytic pathways can, in general, be classed as serine proteases¹, with a catalytic mechanism which operates via acyl-enzyme intermediates². Chase and Shaw³ showed that *p*-nitrophenyl-*p*'-guanidinobenzoate could specifically acylate the active centre of trypsin-like enzymes, giving rise to a stable *p*-guanidinobenzoyl enzyme and other stable acyl-enzymes have since been described^{4,5}. We report here the fibrinolytic use of acylated derivatives of plasmin (E.C.3.4.21.7) and streptokinase-plasmin(ogen) complexes.

An acyl-plasmin, catalytically inert and therefore unable either to degrade systemic plasma proteins or react irreversibly with its principal plasma inhibitors (α 2-antiplasmin^{6,7} and α 2-macroglobulin⁸), may still bind to a fibrin clot because the lysine/fibrin binding sites on the 'kringle' domains of plasmin are structurally and functionally separate from the catalytic centre⁹ (Fig. 1). Deacylation may then occur to give fibrin-bound plasmin. Wiman and Collen¹⁰ have shown that the involvement of the lysine binding sites of plasmin in binding to fibrin reduces the effectiveness of α 2-antiplasmin and provides part of the molecular basis for selective fibrinolysis *in vivo*¹¹. Thus, if an acyl-plasmin is administered to an animal, the release of plasmin by deacylation of fibrin-bound acyl-plasmin would be expected (within certain kinetic limits) to lead to fibrinolysis. In the systemic circulation, however, α 2-antiplasmin would be expected to act rapidly to neutralize the plasmin released by deacylation, provided that the rate and dose were not so great as to consume the α 2-antiplasmin content of the plasma. The concept of active centre acylation may also be applied to plasminogen activators, provided that they too possess fibrin affinity independent of the catalytic centre. In that case, acylation would allow the activator to reduce interactions with plasma plasminogen which lead, during current therapy in man, to harmful plasminemia¹². The putative advantages of therapeutic thrombolysis with acyl-plasmins and acyl-activators, therefore, are (1) enhanced efficacy, due to evasion of inhibitor systems; (2) sustained-release pharmacokinetics, leading to simplified administration and easier clinical control; (3) reduction of

systemic toxicity attributable to hyperplasmaemia; and (4) protection of the enzymes against autolytic degradation.

A general method for the specific synthesis of acyl-enzymes was developed using *p*-aminodiphenyl esters which function as 'inverse' acylating agents¹³. Acyl group radiolabelling demonstrated that active-site acylation with these reagents was quantitative in conditions where nonspecific protein modification was negligible. We prepared acyl derivatives of both human and porcine plasmins, streptokinase-human plasmin activator complex¹⁴ and urokinase, and studied their deacylation rates¹⁵. Using *p*-³H-methoxybenzoyl human plasmin, we found that plasma components on fibrin did not affect the deacylation rate of the acyl-enzyme, and we have confirmed (using immobilized fibrin monomer¹⁶) that active site acylation of plasmin did not affect its affinity for fibrin. Experiments in which *p*-anisoyl human plasmin was mixed with partially purified human $\alpha 2$ -antiplasmin⁷, incubated, re-isolated by affinity absorption on L-lysine-Sepharose¹⁷ and deacylated, showed that an acyl-plasmin did not react irreversibly with the inhibitor. Also, delayed-action fibrinolysis could be demonstrated *in vitro* by incubating human plasma clots with acyl-plasmin, washing and reincubating for a sufficient time to allow deacylation of the absorbed enzyme. Net lysis with acyl-enzymes was significantly greater than lysis by unmodified plasmins.

The kinetics of $\alpha 2$ -antiplasmin action are sufficiently understood¹⁰ to allow the calculation of the steady-state concentrations of free plasmin corresponding to a given rate of deacylation and a given level of acyl-enzyme. Our experimental data in animals have indicated that a steady-state plasma concentration of plasmin amidolytic activity above ~ 10 nM is associated with fibrinogen depletion. It is possible to calculate the relationship between the deacylation rate and dose of an acyl-plasmin that would be consistent with lack of fibrinogen depletion *in vivo*. An acyl-plasmin, at a single dose that would give an initial plasma concentration somewhat less than the known concentration of $\alpha 2$ -antiplasmin in human plasma ($\sim 0.7 \mu\text{M}$)¹⁰, would require a

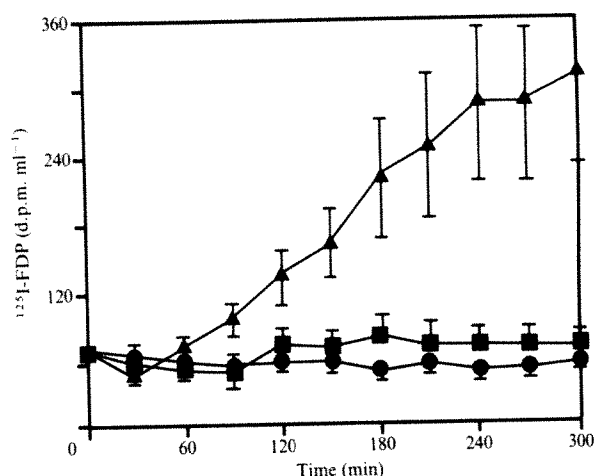


Fig. 2 Release of ^{125}I -fibrin degradation products into the bloodstream of rabbits treated with acylated (▲) and free (■) human plasmins and 0.9% saline solution (●). ^{125}I -fibrin thrombi were prepared in the inferior vena cava of rabbits as detailed elsewhere¹⁵ with the modification that a single strand of 4-ply wool was inserted into the isolated vein segment before thrombus formation to act as a clot anchor. Human plasmin (Kabi batch DtB 23 and Choay batch P299 Pm) was assayed using active site titration³ and amidolytic activity²⁰, and dialysed against 20% v/v glycerol, 0.9% w/v NaCl at 4 °C for 2 h before administration at a dose of 10 mg active enzyme per kg in ~ 20 ml, by constant infusion over 2 h. BRL 26920 was prepared in solution from the above plasmins and 1.0 mM *p*-aminodiphenyl-*p'*-anisate. HCl at 4 °C, pH 7.4 for 1 h, followed by dialysis as above. Complete acylation was checked by amidolytic assay²⁰ and administration was by infusion over 5 min. Control saline, 0.9% w/v, was given over 2 h. ^{125}I -FDP in blood samples was determined by well scintillation counting in a manual bioscint γ counter (Panax) and is given as mean ($\pm$ s.e.m.) concentrations for 12 (BRL 26920), 7 (free plasmin) and 8 (control) animals. Final lysis figures are given in Table 1.

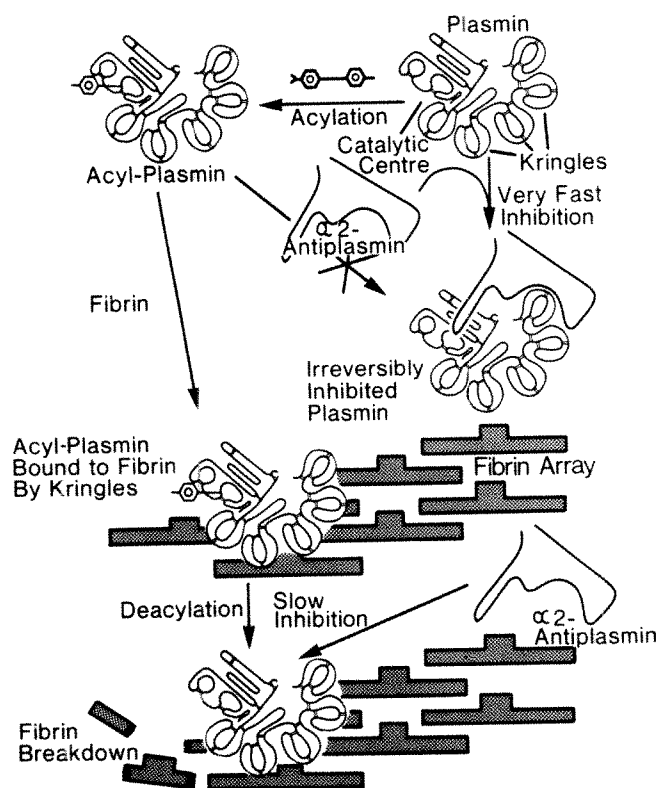


Fig. 1 Schematic outline of the mechanism of fibrinolysis by acyl-plasmins. For description see text. Polypeptide chain structure of plasmin was adapted from ref. 9.

deacylation rate constant of $< 1.0 \times 10^{-3} \text{ s}^{-1}$ if the steady-state level of systemic plasmin were to be kept below 10 nM. One can apply kinetic data on the plasmin- $\alpha 2$ -antiplasmin reaction¹⁸ to a calculation of the steady-state concentration of fibrin-bound plasmin necessary to sustain fibrinolysis in the presence of $\alpha 2$ -antiplasmin, using reasonable estimates of the apparent first-order rate constant for breakage of a single peptide bond in the fibrin-plasmin complex, and an estimate of the number of peptide bonds whose cleavage suffices to detach one fibrin monomer from polymeric fibrin¹⁹. Calculation then suggests that a steady-state fibrin-bound plasmin concentration > 0.1 nM may be necessary to achieve clot breakdown within a few hours. Acyl-plasmins given at a dose of ~ 0.5 – $1.0 \mu\text{M}$ should achieve this steady-state situation if their deacylation rate constants are greater than $\sim 10^{-5} \text{ s}^{-1}$. Acyl-plasminogen activators require a more complex analysis to take into account the relative rates of activation of plasma and fibrin-bound plasminogen. Control of fibrinolysis *in vivo* by manipulation of acyl-activator deacylation rates could be expected to be correspondingly more difficult, depending in practice on factors such as the plasminogen content of the fibrin clot, and the breakdown and clearance mechanism for both acylated and free activator. After analysis of the steady-state kinetics we concluded that deacylation rate constants of $\sim 10^{-4} \text{ s}^{-1}$ should be consistent with thrombolytic activity for both classes of fibrinolytic agent. Two such agents, *p*-anisoyl human plasmin (BRL 26920), with a deacylation half life of 96 min, and a *p*-anisoyl human plasminogen-streptokinase activator (BRL 26921), with a deacylation half life of 40 min (both measured at 37 °C, pH 7.4), have been selected for further evaluation.

We studied the effects of the two agents compared with saline controls, human plasmin and various streptokinase-plasmin complexes in models of venous thrombosis in the rabbit and the

beagle dog (to be published, but partly described elsewhere¹⁵). In both models thromboplastin-induced venous clots radiolabelled with human ¹²⁵I-fibrin were formed around an implanted woollen thread anchored into a major vessel. Lysis was studied by release of ¹²⁵I-fibrin degradation products (¹²⁵I-FDP) into the circulation and by measurement of clot radioactivity at the end of the experiments.

In the rabbit model, acyl-plasmins caused significant thrombolysis when given intravenously as single boluses. Figure 2 shows the effect in rabbits of 10 mg per kg of activator-free free human plasmin (infused over 2 h due to its acute toxicity) and the equivalent bolus dose of BRL 26920, compared with controls given saline. The free plasmin and saline groups showed no significant release of radiolabel into the circulation and gave low, not significantly different, end lysis figures (Table 1). The BRL 26920 group showed sustained release of radioactivity over 5 h and significant net thrombolysis. Free plasmin depleted fibrinogen almost completely, whereas BRL 26920 showed significantly less fibrinogenolysis (Table 1). Plasmin amidolytic activity (chromogenic substrate assay²⁰) in the plasma of animals treated with plasmin was much greater than in animals treated with BRL 26920. The thrombolytic activity of BRL 26920 was also observed in dogs (Table 1).

Because of the known species differences in the mechanism of thrombolysis by streptokinase^{21,22}, BRL 26921 was compared in animals with non-acylated streptokinase-plasmin activator complexes prepared in various conditions (Table 1). Of the unmodified activators, only the minimally degraded form (Type 2) showed significant thrombolytic activity. When degraded activator (Type 1) was *p*-anisoylated at the active centre, the acyl-activator showed increased thrombolytic activity at the equivalent dose. BRL 26921 was tested in groups of rabbits at

three dose levels (Table 1) and showed significant thrombolysis compared with the saline controls. At 2,500 units per kg, all BRL 26921 preparations were consistently more active than all non-acylated activator preparations ($P < 0.01$) and a significant slowing in the rate of ¹²⁵I-FDP release was observed when animals that responded to non-acylated activators were compared with those given acylated activators. None of the tested agents significantly altered fibrinogen levels in the rabbits.

BRL 26921 also showed highly significant thrombolytic activity (measured radiochemically and venographically) against freshly formed thrombi in dogs at two dose levels when compared with saline controls and the least degraded activator form (Table 1). A sustained release of ¹²⁵I-FDP over 4–5 h followed a single dose of BRL 26921. In another group of dogs, the thrombi were aged by isolation from the circulation for 24 h. Isolating ligatures were removed and the dogs treated with BRL 26921 at a single dose of 5,000 units per kg. Radiochemical measurements and venography indicated good thrombolysis (Table 1).

We believe that acyl-enzymes are a novel class of therapeutic agents, in which suitable acyl group modification can provide enzyme precursors whose properties can be manipulated to suit various clinical situations, depending on whether fast or slow thrombolysis is required. Because of their novel pharmacokinetics *in vivo*, they could be expected to show advantages over currently used thrombolytic agents, particularly in producing reduced hyperplasmaemia. We consider that BRL 26921 (which is the more active and accessible agent) may lead to useful improvements in the management of venous thrombosis and some types of arterial thromboembolic disease²³.

We thank Dr J. Cash for advice and help over several years, and also Kabi (Stockholm), Choay (Paris), Novo (Copenhagen)

Table 1 Summary of thrombolytic action of agents in rabbit and dog models of venous thrombosis

Model	Agent	Dose SK units (mg) per kg	No. of animals	% Lysis	% Fibrinogen level	Time to peak ¹²⁵ I-FDP levels (min)
Rabbit	Saline	—	8	3 ± 1	110 ± 1	—
Rabbit	Free human plasmin	(10)	7	7 ± 1	4 ± 3	—
Rabbit	SK-Pm activator Type 1	2,500	5	5 ± 1	95 ± 1	—
Rabbit	SK-Pm activator Type 2	2,500	6	20 ± 4	106 ± 3	49 ± 8
Rabbit	SK-Pm activator Type 3	2,500	5	14 ± 4	—	—
Rabbit	<i>p</i> -Anisoyl activator Type 1	10,000	5	32 ± 5	99 ± 4	36 ± 6
Rabbit	BRL 26921	2,500	5	27 ± 7	112 ± 25	162 ± 35
Rabbit	BRL 26921	2,500	9	31 ± 3	95 ± 2	183 ± 32
Rabbit	BRL 26921	5,000	5	38 ± 10	95 ± 2	132 ± 44
Rabbit	BRL 26921	10,000	5	57 ± 11	91 ± 2	144 ± 40
Rabbit	BRL 26921 (lyophilized)	2,500	8	45 ± 6	—	198 ± 23
Rabbit	BRL 26920	(10)	12	21 ± 4	71 ± 8	≥ 300
Dog (short-term)	Saline	—	5	14 ± 5	98 ± 3	—
Dog (short-term)	SK-Pm activator Type 2	1,500	5	10 ± 4	93 ± 5	—
Dog (short-term)	SK-Pm activator Type 2	2,500	5	25 ± 18	107 ± 8	—
Dog (short-term)	BRL 26921	1,500	5	77 ± 14	105 ± 3	312 ± 20
Dog (short-term)	BRL 26921	2,500	5	85 ± 14	109 ± 5	—
Dog (short-term)	BRL 26920	(10)	5	65 ± 9	91 ± 2	330 ± 13
Dog (long-term)	Saline	—	5	38 ± 7	100 ± 7	—
Dog (long-term)	BRL 26921	5,000	5	98 ± 2	100 ± 8	276 ± 78

Activator complexes were prepared from equimolar proportions of streptokinase (SK, Kabi) and lys-plasminogen (Pm, Kabi, K 2086) at a total protein concentration of 1 mg ml⁻¹ in 0.1 M Tris-HCl, 0.9% w/v NaCl, 20% v/v glycerol, pH 7.4. Type 1 activator (degraded) was prepared at 25 °C for 40 min, Type 2 activator (minimally degraded) at 0 °C for 10 min; Type 3 activator was a commercial lyophilized preparation (Behringwerke, Lot 9018) and was dissolved in 0.9% w/v NaCl at a concentration of 1.4 × 10⁵ U ml⁻¹. Acylation of activator complexes was carried out using 1.0 mM *p*-amidinophenyl *p*'-anisate. HCl at 0 °C for 1 h. Other conditions as above. BRL 26921 was prepared in solution by addition of acylating agent to streptokinase, followed by addition of plasminogen and reaction at 0 °C for 1 h. Lyophilized BRL 26921 was prepared on a large scale in solution, diafiltered exhaustively to remove excess acylating agent and lyophilized in the presence of L-lysine, HCl, D-mannitol and human serum albumin. Doses were based on streptokinase contents of activator complexes and administered in 2.0 ml of 0.9% w/v NaCl as a single bolus over 2 min. For plasmin and BRL 26920 dosage and preparation, see Fig. 2 legend. Fibrinogen level was estimated by a modification of the spectrophotometric method of Burmester *et al.*²⁴ at 280 nm. Percentages of the initial level averaged over the experimental period (5 h for rabbits, 6 h for short-term dogs) and over the experimental group. % Lysis, % fibrinogen level and time to peak ¹²⁵I-FDP levels are all means (±s.e.m.). For rabbit model see ref. 15 and Fig. 2. The short-term dog model was similar in outline to the rabbit model, except that thrombi were located in a section of the femoral vein. Homogeneous radiolabelling of blood before clotting around a woollen thread was achieved by aspirating the blood within an isolated vein segment through a collateral vein into a syringe containing radiolabelled fibrinogen. Blood was then returned to the evacuated segment which had been injected with rabbit brain thromboplastin, and restricting ligatures removed after clotting. In the long-term modification of the dog model, restricting ligatures were kept in place to isolate the thrombus from the circulation for 24 h. Treatment was initiated after removal of the ligatures.

and Behringwerke (Marburg) for generous gifts of materials. We also thank Mr I. Dodd, Mr C. Beilby, Miss H. C. Grundy and, in particular, Mr M. J. Hibbs for technical assistance.

Received 26 September 1980; accepted 21 January 1981.

1. Griffin, J. H. *et al.* *Meth. Enzym.* **45**, 56–177 (1976).
2. Bender, M. L. & Zerner, B. *J. Am. chem. Soc.* **84**, 2550–2555 (1962).
3. Chase, T. Jr & Shaw, E. *Biochem. biophys. Res. Commun.* **29**, 508–514 (1967).
4. Chase, T. Jr & Shaw, E. *Biochemistry* **8**, 2212–2224 (1969).
5. Glover, G., Wang, C.-C. & Shaw, E. *J. med. Chem.* **16**, 262–267 (1973).
6. Moroi, M. & Aoki, N. *J. biol. Chem.* **251**, 5956–5965 (1976).
7. Mullertz, S. & Clemmensen, I. *Biochem. J.* **159**, 545–553 (1976).
8. Barrett, A. J. & Starkey, P. M. *Biochem. J.* **133**, 709–724 (1973).
9. Sottrup-Jensen, L. *et al.* in *Progress in Chemical Fibrinolysis and Thrombolysis* Vol. 3 (eds Davidson, J. F. *et al.*) 191–209 (Raven, New York, 1978).
10. Wiman, B. & Collen, D. *Eur. J. Biochem.* **84**, 573–578 (1978).
11. Wiman, B. & Collen, D. *Nature* **272**, 549–550 (1978).
12. Brogden, R. N., Speight, T. M. & Avery, G. S. *Drugs* **5**, 357–445 (1973).
13. Tanizawa, K., Kasaba, Y. & Kanaoka, Y. *J. Am. chem. Soc.* **99**, 4495–4499 (1977).
14. McClintock, D. K. & Bell, P. H. *Biochem. biophys. Res. Commun.* **43**, 694–702 (1971).
15. Smith, R. A. G. *Eur. Patent Applic.* No. 9879 (1980).
16. Heene, D. K. & Matthias, F. R. *Thromb. Res.* **2**, 137–154 (1973).
17. Deutsch, D. D. & Mertz, E. T. *Science* **170**, 1095–1096 (1970).
18. Wiman, B. & Collen, D. in *The Physiological Inhibitors of Coagulation and Fibrinolysis* (eds Collen, D. *et al.*) 145–146 (Elsevier, Amsterdam, 1979).
19. Mihalyi, E. *et al.* *Biochemistry* **15**, 5372–5381 (1976).
20. Friberger, G. *et al.* in *Chromogenic Peptide Substrates* (eds Scully, M. F. & Kakkar, V. V.) 121–127 (Churchill-Livingstone, Edinburgh, 1979).
21. Klöcking, H.-P. in *Fibrinolysis and Antifibrinolysis* (ed. Markwardt, F.) 151–153 (Springer, Berlin, 1978).
22. Baillie, A. J. & Sim, A. K. *Throm. Diath. haemorrh.* **25**, 499–506 (1971).
23. Aber, C. P. *et al.* *Br. med. J.* **2**, 1100–1104 (1976).
24. Eurmester, H. B. C. *et al.* *J. clin. Path.* **23**, 43–46 (1970).

Supposed lamarckian inheritance of immunological tolerance

L. Brent*, L. S. Rayfield*, P. Chandler†, W. Fierz†, P. B. Medawar† & E. Simpson†

* Department of Immunology, St Mary's Hospital Medical School, London W2 1PG, UK

† Transplantation Biology Section, Clinical Research Centre, Harrow, Middlesex HA1 3UJ, UK

The notion that an adaptation acquired during an organism's lifetime can somehow be imprinted on the genome and so become heritable has been faulted by every critical test to which it has hitherto been exposed, but many naturalists have lost their faith in what seems to them to be the all-encompassing explanatory glibness of neo-darwinism. Although this criticism is unfair it is entirely proper that neo-darwinism should be under constant critical scrutiny. Interest was therefore aroused by the claim of Gorczynski and Steele¹ that tolerance of A strain antigens induced in CBA mice by injecting into them (CBA × A/J)F₁ spleen and bone marrow cells could be transmitted down the male line. Such a claim is of particular interest because spermatozoa, unlike oocytes, are produced throughout life and might thus conceivably have participated in the mechanism envisaged by Steele² as that responsible for the supposed transfer of genetic information from soma to germ plasm. It was claimed that as many as 60% of the progeny of tolerant fathers mated with normal CBA females were unresponsive or hyporesponsive in an *in vitro* assay in which their spleen cells were tested for reactivity against A/J strain histocompatibility antigens in the cell-mediated lympholysis assay (CML). We describe here our failure to confirm these findings and our inability to extend them by testing the progeny for their reactivity against skin allografts from the tolerance-conferring strain.

Newborn CBA/Ca (H-2^k) mice (CBA) from six litters were made tolerant of A/J (H-2^d) (A) histocompatibility antigens by the intraperitoneal inoculation of 5×10^7 adult male (CBA × A)F₁ hybrid lymphoid cells, consisting of a 1:10 mixture of bone

marrow and spleen cells. Ten males received the same dose of spleen cells every 2 weeks throughout the breeding programme. The remainder, 13 females and 4 males, received skin grafts from male A strain donors 8 weeks after birth: instead of being sloughed off in 10–11 days, grafts were fully accepted and they remain healthy at the time of writing (11 months later). Of the 10 males selected for breeding, 4 (nos 1–4) received A strain skin grafts 1 week after mating; these animals were likewise fully tolerant. The other six males (nos 5–10) did not receive their skin grafts until the completion of breeding 6 months after birth; these too proved to be fully tolerant.

Each tolerant male was mated with three normal CBA females, and the progeny were marked so that each individual was identifiable throughout the experiment, thus making it possible to compare litters from individual fathers in the two assays used (CML and grafting). A similar procedure was followed in the production of progeny from normal CBA fathers. A total of 201 experimental progeny and 138 normal (or control) progeny were raised; the immunological reactivity of the majority was assessed using the CML assay, and almost all received strain A or (CBA × A)F₁ hybrid skin allografts after completion of the *in vitro* tests. Although Gorczynski and Steele had used only splenic lymphocytes for CML testing, peripheral blood lymphocytes (PBLs) were used by us for a large number of mice because this allowed repeated observations with minimal interference and the transplantation of skin grafts to the same animals. Splenic lymphocytes were used in other experiments; these were obtained by removing a small portion of the spleen, thus ensuring that these animals had splenic tissue at the time of skin grafting. About two-thirds of all hemisplenectomized mice had previously been tested at least once using PBLs. The latter were prepared from heparinized blood drawn from a tail vein, hemisplenectomy was carried out under full Sagatal anaesthesia through a small lateral incision, and skin grafting was done according to the technique of Billingham and Medawar³. Skin grafts were first inspected on day 10 or 11 and thereafter scored daily, necrosis of the whole graft surface providing the end point. Although strain A tail skin was used in the earlier experiments, most of the progeny received (CBA × A)F₁ hybrid grafts because it was felt that this might favour longer survival in hyporeactive animals, a point to which Steele drew our attention.

The *in vitro* sensitization and ⁵¹Cr release CML assay were performed on PBLs as described by Chandler *et al.*⁴. Briefly, PBLs were separated from red cells by hypotonic shock and suspended at 1×10^6 per ml in bicarbonate-buffered RPMI medium supplemented with fetal calf serum, 10 mM HEPES, 5×10^{-5} M 2-mercaptoethanol, glutamine, penicillin and streptomycin. Aliquots (100 µl) of this suspension were placed in each well of round-bottomed microtitre plates (Sterilin M24

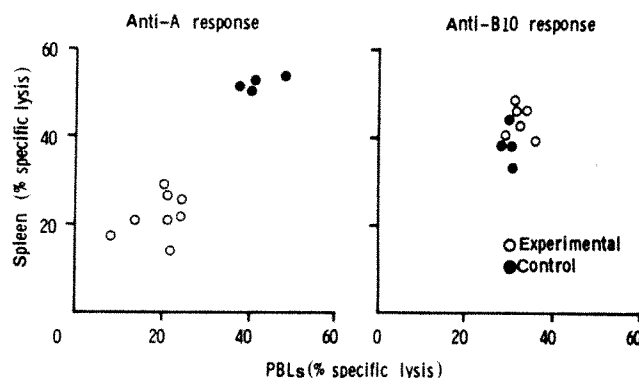


Fig. 1 Cytotoxic responses of experimental and control fathers (spleen versus PBLs). Cytotoxicity is at a K:T ratio of 5:1, derived from regression analysis.

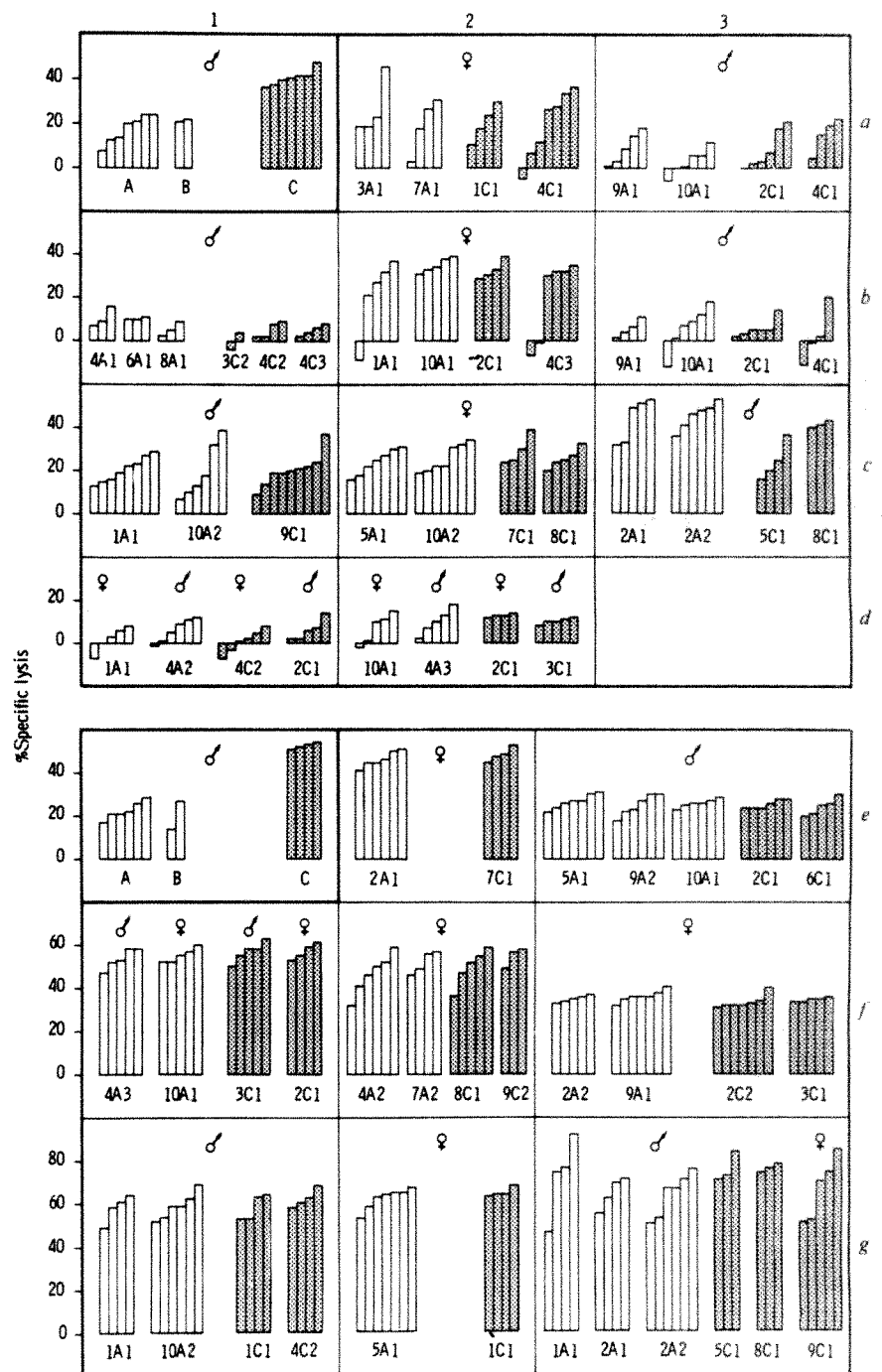


Fig. 2 Anti-A cytotoxic responses of PBLs (panels a–d) and spleen cells (e–g). A, experimental fathers, also shown in Fig. 1. B, males given only one neonatal inoculation of 5×10^7 (CBA $\times$ A) F_1 spleen and bone marrow cells. C, control fathers. The progeny are identified by the number of the father followed by A or C and then the individual litter number. Each panel represents the data from a separate experiment, with sex of the responders indicated. Cytotoxicity is at a K:T ratio of 5:1, derived from regression analysis.

ART) and the same volume and concentration of irradiated (2,000 rad) allogeneic spleen cells were added as antigen. Three to six wells of each PBL sample were set up against the allo-antigen being tested, A/J, and whenever PBL recovery had been sufficient (in over 70% of samples), a further three to six wells were set up against 'third party' C57BL/10 alloantigen. After 5 days' incubation at 37 °C and in a humidified 5% CO₂ atmosphere, the responder ('killer') cells were collected and the appropriate wells for each sample pooled, counted and dispensed in four twofold dilutions into fresh microtitre plates. 5×10^3 or 1×10^4 ^{51}Cr -labelled 24-h concanavalin A blasts from spleen cultures of A/J or C57BL/10 were then added. The highest killer: target ratio (K:T) was 10:1. The plates were spun briefly and incubated for 3 h before removal of an aliquot of the supernatant for γ -counting. Percentage specific lysis was taken

as $(E - C/M) \times 100$, where E was the c.p.m. for the experimental supernatant, C the c.p.m. for the medium control supernatant and M the c.p.m. of target cells lysed with 5% Triton. Regression analysis of the percentages of specific lysis titrated against K:T ratio was performed and from this curve the percentage of specific lysis at a ratio of 5:1 was obtained. The r^2 value (r = the coefficient of correlation) of each curve was recorded and used in the assessment of the results.

Spleen cells from partial splenectomy were set up in culture in the same RPMI-supplemented medium as the PBLs, but at a concentration of $5 \times 10^6 \text{ ml}^{-1}$. Aliquots (2 ml) of this suspension were incubated in 7-ml Linbro wells (Flow Laboratories 76-053-05) together with 2 ml of irradiated spleen cells ($5 \times 10^6 \text{ ml}^{-1}$) as antigen. The cells were cultured for 5 days and the assay performed as for PBLs, except that the ^{51}Cr -labelled

Table 1 Survival of skin allografts transplanted to experimental and control progeny

Father	Sex	Experimental progeny				Control progeny			
		No.	No CML†	PBL‡	PBLs/spleen§	No.	No CML†	PBL‡	PBLs/spleen§
1	♂	10	—	11 13	<11 <11 <11 <11 <11 12½ 13½ 15½	3	—	—	<11 <11 13½
	♀	5	—	—	<10 <10 10 12 12	3	—	—	<11 <11 <11
2	♂	12	<11 16½	—	<11 <11 <11 <11 <11 12½ 12½ 12½ 13½ 14	13	<10 <10	—	<10 10 <11 <11 <11 11½ 12½ 12½ 13 13½ 16½
	♀	14	<11	—	<11 <11 <11 <11 <11 <11 <11 <11 <11 <11 <11 12½ 13½	20	—	<11 <11 <11 <11 <11 <11 <11 <11 <11 12½	<11 <11 <11 <11 <11 <11 <11 <11 11½ 14½
3	♂	6	12½ 15½ 15½	11½ 13 14½	—	9	<10 13	<11 17½	<11 <11 <11 11½ 11½
	♀	6	<11 <11	<11* <11* <11* <11*	—	6	—	<11 <11	<11 <11 <11 <11
4	♂	20	<11	10½ <11* <11* <11* <11* 11 11½ 13½*	<10 <10 <10 <10 <11 <11 <11 <11 14 14½ 14½	16	—	<11 <11 <11 <11 11½ 12½ 13½ 14½	<11 <11 <11 <11 <11 11 11½ 15½
	♀	12	<11	<11* <11* <11* <11* <11*	<11 <11 <11 <11 <11 12½	13	—	<11 <11 <11 <11 <11 <11 12½	<10 <10 <10 <10 <10 11
5	♂	14	11 15½	<11* <11* <11* <11* <11*	<11 <11 <11 13½ 14½ 15½ 16½	3	—	—	<11 <11 <11
	♀	9	—	<11* <11* <11* <11*	<11 <11 <11 12½ 14½	2	<10 <10	—	—
6	♂	7	<11 <11 11½	11½ 11½ 11½ 12½	—	8	10 14½ 14½	—	<11 <11 <11 13½ 15½
	♀	0	—	—	—	1	<10	—	—
7	♂	16	<10 <11 11½ 12½ 12½ 13 13½ 13½ 13½ 14½ 14½	11½ 11½ 12 12½ 14½	—	2	<10 11	—	—
	♀	8	—	<11* <11* <11* <11*	<11 <11 <11 <11	4	—	—	<11 <11 <11 <11
8	♂	6	<11 <11 11½	11½ 11½ 15½	—	2	—	—	<11 <11
	♀	7	<10 <10 <10 11½	<10 10 14½	—	5	—	—	<11 <11 <11 <11 <11
9	♂	5	—	—	<11 <11 <11 13½ 14½	9	<11 <11 <11 12	<11 13 14 14½ 15½	—
	♀	9	<11 <11	—	<11 <11 <11 <11 <11 <11 14½	7	—	—	<11 <11 <11 <11 <11 <11 <11
10	♂	12	—	—	<11 <11 <11 <11 <11 <11 <11 <11 13½ 13½ 14½ 15½	2	13½ 13½	—	—
	♀	15	<10 11½ 11½	<11 <11 <11 <11 <11 13½ 17	<11 <11 <11 <11 <11	2	<10 <10	—	—
Totals	♂	108	25	30	53	67	15	15	37
	♀	85	13	27	45	63	5	19	39
Median survival times of	♂		11½	11½	<11		<11	<11	<11
	♀		<11	<11	<11		<10	<11	<11

Survival times are for (CBA×A)_F₁ hybrid grafts except those marked by an asterisk, which were derived from A strain donors.

† Progeny not tested in CML.

‡ Only tested with PBLs.

§ Tested with PBLs and spleen cells.

target cells were increased to 1×10^5 per well; the highest K:T ratio was again $\sim 10:1$.

Both PBL and spleen samples were scrupulously randomized and coded at the time of their removal from the mice. This code was not broken until after the CML results had been computed.

In most experiments, responders of only one sex were used. Spleen cells from female mice were used for antigen and as targets whenever the responder cell donors were females. Male mice were normally stimulated by, and tested on, male cells but

on a few occasions (and always in experiments where both males and females were used as responders) female antigen and target cells were used. In this way the analysis remained uncomplicated by anti-H-Y reactivity.

The results of testing the tolerant (designated A) and the normal (designated C) fathers in the CML test, using both PBLs and spleen, are shown in Fig. 1 and in Fig. 2, panels 1a and 1e. The two groups are clearly distinguishable from each other by their CML reactivity to A strain, but not to third party B10,

alloantigens. The tolerant CBA fathers were thus specifically hyporesponsive to A strain antigens and this hyporesponsiveness is shown equally by splenic and peripheral blood lymphocytes. Nevertheless, they were considerably more responsive than were the tolerant CBA fathers in Gorczynski and Steele's experiments¹. The reason for this difference cannot be that in our study CML reactivity was tested at a rather late stage (~10 months after birth) because this was also true of the experimental fathers tested by Gorczynski and Steele¹ (8–10 months). It could, however, reflect the fact that the control CBA-anti-A/J response in Gorczynski and Steele's experiments was substantially lower than that in many of our experiments or that, contrary to our practice, their assessment was made with control mice that had not been age-matched. Bearing in mind the advanced age of the experimental fathers, this makes a direct comparison of paternal CML responsiveness in the two studies difficult.

The CML data of the progeny (Fig. 2, panels 2a–3d and 2e–3g) do not show any consistent pattern of hyporeactivity towards A alloantigens. These data include all spleen and PBL experiments which met the technical criteria of (1) the spontaneous medium release of ⁵¹Cr by target cells being less than 40% of maximum release and (2) at least 60% of the responders showing cytotoxic activity which titrated and which, on regression analysis, had an r^2 value of 0.80–1.00. The number of mice tested and assays performed that met these criteria are shown in Fig. 3, which also provides the pooled anti-A results of (1) all experiments performed using PBLs, (2) the data selected from the PBL experiments according to the criteria above and (3) spleen data, all of which met the criteria with the exception of one experiment. It is clear from Fig. 3a–c that there is no essential difference between the response of progeny of tolerant and of control fathers.

However, the pooling of data from many different experiments, each using a different batch of target cells, is not acceptable for statistical analysis, and weight was therefore placed only on the comparisons of animals tested within a single experiment. The data of 20 such experiments are shown as individual panels in Fig. 2. The results of each have been analysed by the Wilcoxon rank order test to determine whether the tolerant fathers or their progeny are significantly different from those of the normal fathers or their progeny. There is a highly significant difference between the response of the tolerant fathers and the control fathers (PBLs, panel 1a, $P < 0.001$; spleen, panel 1e, $P < 0.005$). There is also a difference in the response of experimental offspring in panel 3c (PBLs), with the male progeny of father 2A giving higher responses than the male progeny of control father 5C plus 8C. This difference is significant at the 0.05% level. However, panel 3g shows that the spleen responses of the same experimental mice were significantly lower ($P < 0.01$) than those of the controls 5C plus 8C, although the differences are quantitatively small. In the spleen experiment, moreover, the anti-B10 responses of the 2A progeny were lower than those of the controls 5C and 8C ($P < 0.02$). One further PBL experiment (Fig. 2, panel 3a) shows a significantly lower response ($P < 0.02$) by some of the progeny of father 10A than by some offspring of control father 4C (but not 2C), although this result was not reproducible (see panel 3b). These two positive results cannot therefore be held to carry any conviction—they seem to be non-reproducible and/or nonspecific. The results given in the remaining panels show no significant difference between the responses of experimental and control progeny. In each experiment the progeny of individual experimental fathers were compared with those of individual control fathers tested on the same occasion. All experiments completed but not shown in Fig. 2 because they failed to meet the technical criteria were nevertheless similarly analysed, again without detectable differences. Table 1 shows the numbers of offspring of each sex tested for each of the experimental and control fathers.

Our investigation extended the scope of the analysis described by Gorczynski and Steele¹ in that almost all offspring, including some that could not be included in the CML assay, eventually

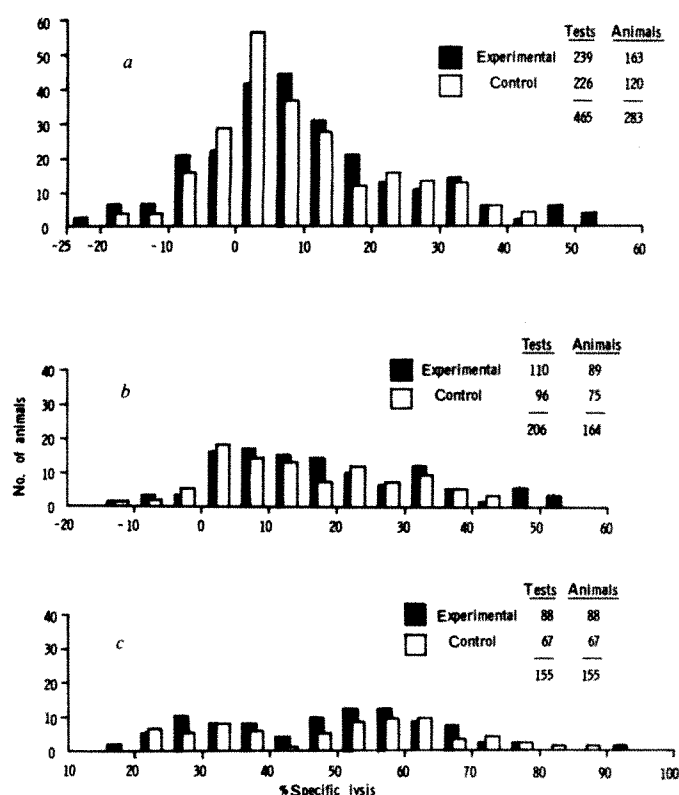


Fig. 3 Anti-A cytotoxic responses of PBLs and spleen-pooled data. *a*, PBLs, unselected data. *b*, PBLs, selected data, $r^2 > 0.80$ in $> 60\%$ of mice in each experiment and spontaneous release $< 40\%$. *c*, Spleen data excluding one experiment. Cytotoxicity is at a K:T ratio of 5:1, derived from regression analysis.

received skin allografts bearing A strain antigens—a total of 193 experimental and 130 control mice. Of these, 80% had been tested for CML reactivity. In addition, the 10 experimental fathers were grafted with A strain skin, 4 of them (nos 1–4) at commencement and the remainder (nos 5–10) after completion of the breeding programme; all these fathers were fully tolerant and carried healthy grafts throughout their subsequent life (up to 11 months). Table 1 gives a detailed breakdown of the individual survival times of grafts transplanted to the experimental and control progeny and it identifies those mice that had been tested in CML and those (the minority) that had not. There is no discernible difference between experimental and control progeny. It is evident that strain A grafts were rejected swiftly and consistently by both male and female recipients, whereas F_1 grafts were rejected more variably, especially by males. The range of survival of grafts on experimental and control mice was closely similar and no graft showed long-term survival.

It is therefore clear that, using the same strain combination, we have been unable to confirm the CML results of Gorczynski and Steele, and the *in vitro* data receive full support from the skin graft assay. Reasons for the discrepancy between the CML data from the two studies can be looked for in the following points.

(1) The initial tolerance-conferring inoculum in our experiments consisted of 5×10^7 cells of a mixture of F_1 hybrid spleen and bone marrow cells, half the number reported by Gorczynski and Steele¹. Furthermore, the ratio of bone marrow cells to spleen cells was 1:10, compared with 1:1 in their study. However, a single dose of 5×10^7 lymphoid cells is certainly adequate for the induction and maintenance of skin graft tolerance and CML hyporesponsiveness (see above and Fig. 2, group B; also ref. 5); indeed, males receiving a single dose at birth had a CML reactivity indistinguishable from that of repeatedly injected males (group B versus group A of Fig. 2, panels 1a and 1e). We would therefore not expect these

differences in protocol to explain the discrepancy if tolerance is the criterion chosen. We have nevertheless begun further experiments in which mice have received 5×10^7 F₁ hybrid spleen and 5×10^7 bone marrow cells at birth, followed by subsequent doses of 5×10^7 spleen and bone marrow cells at a ratio of 1:1. Data from this additional study will not be available for some time.

(2) Although Gorczynski and Steele provide some documentation on the number of experimental fathers and their progeny, there is no information about their controls, which were presumably taken from standard stock without reference to known fathers. This point has been taken care of in our study.

(3) Although the sex of the experimental progeny is clearly shown in Gorczynski and Steele's pedigree diagram, there is no information about the sex of the controls in any one experiment (series), nor is it possible to ascertain what the sex of the stimulator and target cells was. It might therefore be argued that the responsiveness of males and females was different and that their representation in control and experimental groups was uneven, and/or that anti-H-Y responses biased their results. We investigated whether male and female responsiveness in the CML assay is indeed different. Two experiments using PBLs are shown in Fig. 2, panels 1d and 2d; there was no significant sex difference. It is true that another experiment using spleen cells (Fig. 2, panels 1g, 2g) did show a statistically significant difference ($P < 0.05$), with males giving a lower response than females, but quantitatively the difference is so small that it is unlikely to account for the discrepancy. Furthermore, as CBA females rarely make primary anti-H-Y responses⁶ their contribution in the Gorczynski and Steele study¹ can be expected to have been negligible.

(4) Gorczynski and Steele did not test either the injected fathers or their offspring with A/J skin allografts. In our study the experimental fathers were found to be fully non-reactive to skin allografts whereas all their offspring uniformly rejected their grafts.

(5) The CML sensitization and effector assays used by us were not precisely the same as those used by Gorczynski and Steele. They tested only spleen cells and the levels of cytotoxicity they reported¹ were much lower than those obtained in our spleen assays, but it is difficult to make comparisons as their report does not make it clear whether precise cell counts and full titrations were performed on the effector cells generated in culture. It is, however, striking that the low level of cytotoxicity described by them is comparable to the low levels observed by us in some of the experiments in which PBLs were used as responders. The cytotoxicity of both control and experimental mice here tended to be more variable and, we therefore reason, more likely to be sensitive to weak suppressor influences. Our PBL results are nevertheless quite indistinguishable from our spleen results, for neither differentiate between experimental and control offspring.

(6) Whereas all mice in our study were derived from strains bred at St Mary's Hospital Medical School, where the homozygosity of inbred strains is checked from time to time by syngeneic skin transplantation, the mice used by Gorczynski and Steele, although derived from the same foundation stock, were obtained from different breeding colonies. It is conceivable that this introduced a measure of genetic variability that could account for small differences in responsiveness.

(7) We used essentially the same kind of statistical analysis as Gorczynski and Steele. This analysis could fail to reveal a small proportion of non- or hyporesponders in individual experiments, but visual examination of the data suggests that this has not been the case.

It is difficult to account for the discrepancy in the two sets of findings by examining differences in the protocols used. There are available data from several other investigations designed to test the hypothesis put forward by Gorczynski and Steele¹, who have recently extended their observations to another strain combination and claim to have shown the simultaneous and independent inheritance of tolerance to two distinct haplo-

types⁷. R. B. Ashman, C. E. Woodhams and R. V. Blanden (personal communication) tested by CML and skin grafting the progeny of male CBA mice made tolerant at birth by injection of (CBA $\times$ BALB/c)F₁ cells; they did not differ from control progeny although one graft survived on an experimental mouse for a surprisingly long time and this mouse gave a low CML response. McLaren *et al.* (see accompanying letter⁸) examined by CML the reactivity of homozygous progeny of male B6 $\leftrightarrow$ BALB/c tetraparental chimaeric mice bred with both B6 and BALB/c females. The B6 progeny were not hyporesponsive to BALB/c alloantigens, nor were the BALB/c progeny hyporesponsive to B6. J. Craig Gray (personal communication) has been assessing the skin graft survival times in offspring of tolerant male rats, using grafts from the cell donor strain, and here again no hyporesponsiveness has yet been uncovered.

Using a rather different approach, Hašek *et al.*⁹ have been studying CML reactivity of first generation backcross mice of F₁ hybrid H-2 heterozygous fathers. It could be argued that it should follow from the hypothesis under discussion^{1,2} that the natural tolerance of an F₁ father is passed on to his homozygous progeny with respect to their reactivity to the alloantigens of his other parental haplotype. This is evidently not so. Along similar lines, and following the much earlier report of Guttman *et al.*¹⁰, Brent and Rayfield (unpublished data) have been breeding CBA fathers tolerant of A with (CBA $\times$ A)F₁ females to examine the ability of the progeny to reject A skin grafts. In this case positive results would be indicated by an abnormal number of offspring bearing long-surviving A strain grafts compared with control offspring of normal CBA males bred with (CBA $\times$ A)F₁ females. Although the study is not yet complete there are, to date, no differences between the control and experimental groups. Moving away from tolerance to alloantigens, A. Müllbacher, W. Fierz and J. Sheena (personal communication) have bred from CBA mice made tolerant to Bebaru virus by neonatal and subsequent weekly injections of virus inactivated by X-ray irradiation. The offspring of some of the tolerant fathers gave low responses to Bebaru virus, but these were not found in every experiment.

These experiments may be regarded as a test case for the validity of Lamarckian inheritance, which relates to the inheritance of acquired adaptations, not to the inheritance of the passively acquired somatic changes. An immunological response is surely the paradigm of physiological adaptations—witness the complete confidence with which an immunologist assumes that a rabbit not yet born will be able to make antibodies against a chemical not yet synthesized. There are technical difficulties in devising critical experiments on the inheritance of antibody formation but these do not apply to specific suppression of immune responsiveness (tolerance): it was therefore enterprising of Gorczynski and Steele to use acquired tolerance as the adaptive response on which to base their examination of Lamarckism. Unfortunately, our own findings do not support theirs and confirm the original observations by Billingham *et al.*¹¹ in the identical strain combination. We still believe that all experiments executed hitherto to corroborate the Lamarckian interpretation¹² can be faulted.

We thank Miss S. Kalsi, Miss M. Gaffey and Mr R. Butler for their technical assistance and the MRC and Joint Standing Research Committee of St Mary's Hospital for their support.

Received 4 February; accepted 9 March 1981.

- Gorczynski, R. M. & Steele, E. J. *Proc. natn. Acad. USA*, **77**, 2871–2875 (1980).
- Steele, E. J. *Somatic Selection and Adaptive Evolution. On the Inheritance of Acquired Characters*, 1–91 (Williams & Wallace, Toronto, 1979; Croom Helm, London, 1980).
- Billingham, R. E. & Medawar, P. B. *J. exp. Biol.* **28**, 385–402 (1951).
- Chandler, P., Matsunaga, T., Benjamin, D. & Simpson, E. *J. immun. Meth.* **31**, 341–350 (1979).
- Beverly, P. C. L., Brent, L., Brooks, C., Medawar, P. B. & Simpson, E. *Transplant Proc.* **5**, 679–684 (1973).
- Brenan, M. J. thesis, Univ. London (1981).
- Gorczynski, R. M. & Steele, D. J. *Nature* **289**, 678–681 (1981).
- McLaren, A., Chandler, P., Buehr, M., Fierz, W. & Simpson, E. *Nature* **290**, 513–514 (1981).
- Hašek, M., Holáň, V. & Chutná, J. in *Cellular and Molecular Mechanisms of Immunological Tolerance* (Dekker, New York, in the press).
- Guttman, R. D., Vosika, G. J. & Aust, J. B. *J. natn. Cancer Inst.* **33**, 1–5 (1964).
- Billingham, R. E., Brent, L. & Medawar, P. B. *Phil. Trans. R. Soc. B239*, 357–414 (1956).
- Medawar, P. B. in *The Uniqueness of the Individual*, 79–107 (Methuen, London, 1957).

Immune reactivity of progeny of tetraparental male mice

Anne McLaren*, Phillip Chandler†, Mia Buehr*,
Walter Fierz† & Elizabeth Simpson†

* MRC Mammalian Development Unit, University College London, London WC1E 6BT, UK

† Transplantation Biology Section, Clinical Research Centre, Watford Road, Harrow, Middlesex HA1 3UJ, UK

Steele and Gorczynski¹ have recently suggested that inbred male mice rendered tolerant at birth to the alloantigens of an H-2 plus non-H-2 incompatible inbred strain can transmit this tolerance or hyporeactivity, as measured in a primary anti-H-2 cytotoxic T-cell test *in vitro*, to their progeny, born of normally reactive females syngeneic with the males. As a corollary, it might be expected that the progeny of tetraparental males which are tolerant because from earliest fetal development they are chimaeric with respect to all tissues, including haematopoietic cells and germ cells, might in turn be tolerant to the other set of paternal alloantigens. We have now found, on the contrary, that inbred progeny of one component of a tetraparental male showed heightened responsiveness to the other paternal alloantigens.

We examined homozygous progeny of two C57BL/6 ↔ BALB/c male tetraparental chimaeras² with respect to their ability to mount primary anti-H-2 cytotoxic T-cell responses *in vitro* to either C57BL/6 (B6) (H-2^b) or BALB/c (H-2^d) and to a 'third party' strain, CBA (H-2^k). The two chimaeras were mated first with BALB/c females and of the resulting heterozygous (BALB/c × B6)_{F1} (agouti) and homozygous BALB/c (white) progeny, the homozygotes were kept and tested. The same tetraparental chimaeras were also bred with B6 females, the agouti (heterozygous B6 × BALB/c) progeny disposed of and the black (homozygous B6) kept and tested. The BALB/c and B6 mothers were mostly used for one litter only. We examined the cytotoxic T-cell responses of a total of 46 offspring from 13 litters, 7 with BALB/c and 6 with B6 mothers. The size of the litters of BALB/c mothers was not statistically different from those of the B6 mothers. There were 26 BALB/c progeny (13 males and 13 females) all of chimaera 1 and 20 B6 progeny (10 males and 10 females) of which 12 were fathered by chimaera 1 and 8 by chimaera 2. Both fathers were coat colour chimaeras but there is evidence only for number 1 being a germ-cell chimaera (see above). Chimaera 2 died before he could be tested for T-cell cytotoxic responses, but chimaera 1 was tested and found to be tolerant to each of the parental strains B6 and

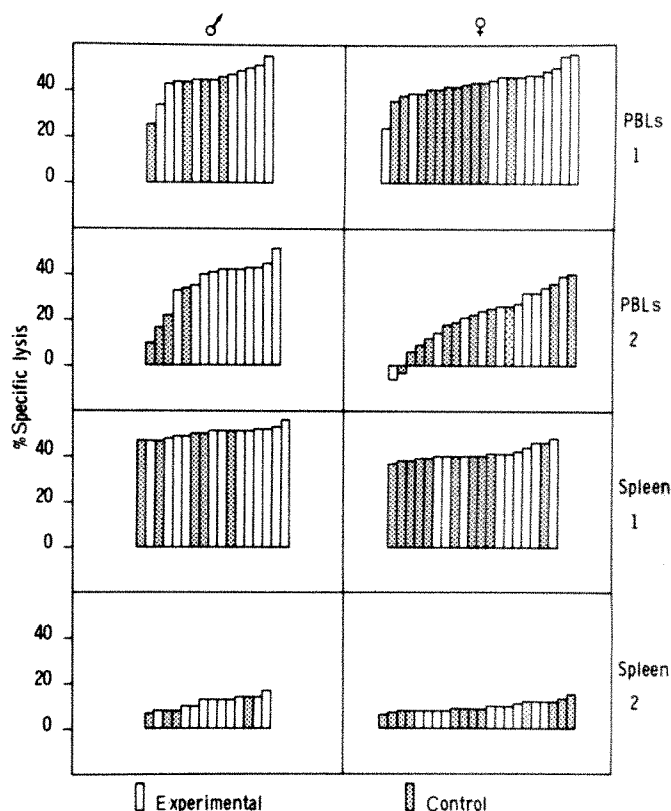


Fig. 1 Anti-C57BL/6 responses of BALB/c progeny of B6 ↔ BALB/c chimaeras and control BALB/c fathers. Conditions of culture. PBLs: two to six 0.2-ml round-bottomed wells were set up with 1×10^5 responders plus 1×10^5 2,000 R irradiated spleen cells per well as antigen, in RPMI with 5% fetal calf serum, 5×10^{-5} M 2-mercaptoethanol, penicillin/streptomycin + glutamine. Spleen: 2×4 ml Linbro wells were set up with 1×10^7 responders + 1×10^7 2,000 R irradiated antigen per well in RPMI medium as above. Assay⁴ was on day 5 of culture. Cells were collected from wells, counted, plated and serially double diluted in round-bottomed 0.2-ml microtitre wells, to which were added a constant number of ⁵¹Cr-labelled target cells (24 h concanavalin A blast cells or thioglycollate-induced peritoneal macrophages) to give attacker:target (A:T) ratios starting at ~10:1. They were collected after 3 h incubation. Experimentally determined % specific lysis was subjected to regression analysis. The height of the bars shown is the % specific lysis at A:T 5:1 taken from this curve ($r^2 > 0.9$). The bars representing individual animals are placed in a ranked order.

BALB/c, but reactive to a third party strain, CBA (H-2^k) (Table 1). Previous investigations of chimaeras made in this laboratory have shown that the inability of such mice to make anti-parental cytotoxic responses does not depend on the level of haematopoietic chimaerism. All chimaeras tested were essentially tolerant³.

We used age-matched BALB/c and B6 litters as controls, born of individually identified fathers (four B6 and six BALB/c) and reared in the same environmental conditions as the chimaera progeny. In each experiment either all males or all females were used as responder, antigen and target. Each control and experimental mouse was tested on four different occasions each 2 weeks apart. On two occasions peripheral blood lymphocytes (PBLs) were prepared from blood samples taken by tail bleeding and set up in micromixed lymphocyte culture (MLC)⁴; on the third occasion a sample of spleen was removed for MLC from each mouse under general anaesthesia; finally, the mouse was killed and the remainder of the spleen taken for MLC. During the experiments, a few mice died and the number in each experiment was further reduced by occasional contamination of a culture.

Table 1 Tolerance to parental antigens of a tetraparental chimaera B6 ↔ BALB/c

Responder	Lysis on target		
	B6	CBA	BALB/c
BALB/c (6)	42	44	ND
B6 ↔ BALB/c	4	40	2
B6 (4)	ND	29	21

Tolerance is expressed as % specific lysis at A:T (attacker:target cell ratio) 5:1 in a cytotoxic T-cell response (method as in Fig. 1). Mean values for BALB/c + B6. ND, not done. Numbers of mice tested are given in parentheses.

Table 2 Mean cytotoxic responses of homozygous progeny of two C57BL/6 $\leftrightarrow$ BALB/c made tetraparental chimaeras

		PBLs		Spleen	
		1	2	1	2
B6 anti-BALB/c	E	40.9 (9)	44.2 (9)	26.4 (7)	11.1 (7)
	C	39.8 (10)	42.9 (10)	29.5 (10)	9.3 (10)
B6 anti-BALB/c	E	13.9* (8)	19.8 (8)	31.7 (8)	10.0 (5)
	C	9.1 (10)	18.8 (10)	31.6 (10)	14.4 (10)
BALB/c anti-B6	E	46.3 (10)	41.5† (11)	50.7 (12)	12.4 (10)
	C	39.9 (4)	20.4 (4)	48.9 (5)	9.2 (4)
BALB/c anti-B6	E	45.8 (11)	24.3 (10)	42.5§ (9)	9.8 (10)
	C	41.5‡ (11)	19.1 (11)	39.8 (10)	9.8 (12)

Mean cytotoxic responses are expressed as % specific lysis at A:T 5:1 in a cytotoxic T-cell response (method as in Fig. 1). E, experimental animals; progeny of B6 $\leftrightarrow$ BALB/c male chimaeras; C, control animals; age-matched normal animals. Numbers in parentheses are the number of animals tested.

* $P < 0.05$; † $P < 0.01$; ‡ $P < 0.01$; § $P < 0.05$ (Wilcoxon rank order test).

The mean anti-BALB/c and anti-B6 cytotoxic responses of the various groups are shown in Table 2. Statistical tests (Wilcoxon rank order test) were done on all experiments. The levels of cytotoxicity generated by B6 chimaera male progeny against BALB/c or third party CBA antigen did not differ significantly from those of age- and sex-matched controls. Among the B6 females cells from the chimaera progeny in one test showed significantly greater ($P < 0.05$) cytotoxicity against BALB/c antigen than did the controls, but this result was not confirmed on the three subsequent occasions, and as the ranking data were statistically heterogeneous, they could not be combined. When the responses of the progeny of chimaeras 1 and 2 were analysed independently the same results were obtained. In contrast, when the response of the BALB/c chimaera progeny to B6 antigen was compared with their BALB/c age- and sex-matched controls, a consistent increase in the level of cytotoxicity of the experimental versus control animals was found. Figure 1 shows the responses of all BALB/c progeny tested: for each sex the rank order results from the four occasions proved to be statistically homogeneous, that is, were consistent, and hence could be combined by ranking the summed rank orders. The excess of experimental over control cytotoxicity levels was, though small, statistically significant ($P < 0.01$) for both the male and female progeny independently. The responses of the BALB/c chimaera progeny to the third party control CBA were no different from that of control animals.

These results are in keeping with the findings of Brent *et al.* (see accompanying letter⁵), in that the state of tolerance manifest in male mice is not transmitted to their progeny. In fact, the difference that we observed was in the opposite direction to that predicted from the work of Gorczynski and Steele¹. It occurred whether or not the BALB/c young had F₁ littermates at birth, so is unlikely to be due to sensitization by antigens of litter mates during pregnancy. We are exploring the possibility that it may be mediated via the mother, perhaps as a result of sensitization by antigens present in the ejaculate.

Received 16 February; accepted 16 March 1981.

1. Gorczynski, R. M. & Steele, E. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2871 (1980).
2. McLaren, A. in *Mammalian Chimeras* (Cambridge University Press, 1976).
3. Matsunaga, T., Simpson, E. & Meo, T. *Transplantation* **30**, 34–39 (1980).
4. Chandler, P., Matsunaga, T., Benjamin, D. & Simpson, E. *J. immun. Meth.* **31**, 341 (1979).
5. Brent, L. *et al. Nature* **290**, 508–512 (1981).

Selective antagonists of benzodiazepines

W. Hunkeler, H. Möhler, L. Pieri, P. Polc, E. P. Bonetti, R. Cumin, R. Schaffner & W. Haefely*

Pharmaceutical Research Department, F. Hoffmann-La Roche & Co. Ltd, CH-4002 Basel, Switzerland

Benzodiazepines produce most, if not all, of their numerous effects on the central nervous system (CNS) primarily by increasing the function of those chemical synapses that use γ -amino butyric acid (GABA) as transmitter^{1,2}. This specific enhancing effect on GABAergic synaptic inhibition is initiated by the interaction of benzodiazepines with membrane proteins of certain central neurones, to which drugs of this chemical class bind with high affinity and specificity^{3,4}. The molecular processes triggered by the interaction of these drugs with central benzodiazepine receptors, and which result in facilitation of GABAergic transmission, are still incompletely understood. Theoretically, benzodiazepines could mimic the effect of hypothetical endogenous ligands for the benzodiazepine receptors, although there is no convincing evidence for their existence⁵; *in vitro* studies indicate that benzodiazepines might compete with a modulatory peptide which is present in the supramolecular assembly formed by GABA receptor, chloride ionophore and benzodiazepine receptor and which reduces the affinity of the GABA receptor for its physiological ligand⁶. The mechanisms of action of benzodiazepines at the molecular level are likely to be better understood following our recent discovery of benzodiazepine derivatives, whose unique pharmacological activity is to prevent or abolish in a highly selective manner at the receptor level all the characteristic centrally mediated effects of active benzodiazepines. Here, we describe the main properties of a representative of this novel class of specific benzodiazepine antagonists.

Working on a series of imidazodiazepines, we found compounds which were potent inhibitors of the specific high-affinity binding of ³H-diazepam to brain synaptosomal fractions, but which failed to produce any of the behavioural and neurological effects typical of benzodiazepines. Detailed analysis revealed that these compounds selectively and potently inhibit the central effects of classic benzodiazepines by acting as their antagonists at benzodiazepine receptors. From a large series of compounds we have selected Ro 15-1788 (Fig. 1) for clinical trials.

Ro 15-1788 is even less toxic than most of the classic benzodiazepines; the following approximate acute LD₅₀ values were determined: 4,000 mg per kg intraperitoneally (i.p.) and 4,300 mg per kg orally (p.o.) in mice, 1,360 mg per kg i.p. and 6,000 mg per kg p.o. in rats. In mice, rats, cats, dogs and squirrel monkeys the compound up to subtoxic doses produces none of the characteristic effects of benzodiazepine drugs, such as sedative, anticonvulsant, muscle relaxant and anticonflict effects. Ro 15-1788 is not a convulsant by itself, it fails to sensitize animals to chemical convulsants such as leptazol and lacks overt central stimulant activity in all species studied (L.P. *et al.*, in preparation).

Specific binding of ³H-diazepam to crude synaptosomal fractions of rat cerebral cortex⁴ was inhibited competitively by Ro 15-1788 (Fig. 2) with an IC₅₀ value of 2 nmol l⁻¹. This is about one quarter the value for diazepam in the same experimental conditions. Ro 15-1788 also dose dependently inhibited the binding in the brain of ³H-flunitrazepam injected intravenously (i.v.) in mice (H.M. *et al.*, in preparation); tested 15 min after

* To whom correspondence should be addressed.

p.o. administration, Ro 15-1788 was about half as potent as diazepam (ED_{50} values 4.0 and 1.7 mg per kg, respectively).

Three test procedures were adapted routinely to screen for potential benzodiazepine antagonism: the anti-leptazol test⁷, the horizontal wire test⁸ in mice and the antagonism of flunitrazepam-induced pre-anaesthesia in squirrel monkeys. In the first test, a supramaximal i.p. dose of leptazol (120 mg per kg) produced tonic seizures in all animals. Diazepam 5 mg per kg i.p. (twice the calculated ED_{90}) fully protected mice from tonic seizures when given 1 h before leptazol. Ro 15-1788 was administered p.o. or parenterally at various doses and varying times after diazepam and the number of animals developing tonic seizures in response to leptazol were counted. When given p.o. 15 min before leptazol (and 45 min after diazepam), Ro 15-1788 antagonized the anticonvulsant effect of diazepam with an ED_{50} of 2.8 mg per kg, but did not reduce the anticonvulsant action of phenobarbitone in the same test procedure.

In the horizontal wire test⁸, central depressant drugs dose dependently impair the ability of mice hung on a wire by their forepaws to heave themselves on to the wire with their hindlegs. Diazepam (3 mg per kg i.p.) reduced the number of mice performing on the horizontal wire by 70–100%. Ro 15-1788 given before, simultaneously with or after diazepam dose dependently attenuated and blocked the depressant effect of diazepam and of other benzodiazepines; when given 15 min after 3 mg per kg diazepam i.p. and tested over 30 min, the ED_{50} value for Ro 15-1788 was 0.2 mg per kg p.o. Ro 15-1788 by itself did not affect the performance of mice and up to very high doses (300 mg per kg) was unable to antagonize the depressant effect of phenobarbitone, meprobamate, ethanol, L-cycloserine (a GABA transaminase inhibitor) and muscimol. However, it antagonized the performance-depressant effect of zopiclone⁹, a non-benzodiazepine with diazepam-like activity and moderate potency as an inhibitor of specific ³H-diazepam binding to brain synaptosomes. A pre-anaesthetic state (loss of righting reflex, ataxia, sleep-like posture) was induced in squirrel monkeys by 3 mg per kg flunitrazepam i.v. When Ro 15-1788 was administered by stomach tube in doses between 10 and 100 mg per kg at the peak of flunitrazepam action, the monkeys stood up spontaneously within minutes and returned to normal behaviour for 2–4 h.

Increase of food-rewarded behavioural responses that are depressed by a punishment contingency is considered to be the most relevant available indication from animal experiments for anxiolytic activity in man. In the procedure used in the present experiments, rats were trained to press a lever for food reward; during 'conflict sessions', in which each lever-press was both rewarded with a food pellet and punished with an electric foot shock, the number of lever presses was markedly reduced. Diazepam (5 mg per kg p.o.) consistently increased the number of punished responses. Ro 15-1788 at 10 mg per kg p.o. prevented the antipunishment effect of diazepam.

Electrophysiological studies revealed the following effects of Ro 15-1788 (P.P. *et al.*, in preparation). In non-anaesthetized acute spinal cats the compound did not itself affect segmental dorsal root potentials and reflexes, polysynaptic ventral root reflexes and the spontaneous discharge rate of γ -motoneurons (for methods see ref. 10). The very potent benzodiazepine, Ro 11-3128 (3-methylclonazepam), at 0.1 mg per kg i.v., markedly

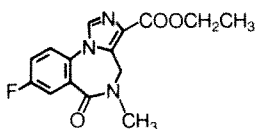


Fig. 1 Structural formula of Ro 15-1788 (ethyl-8-fluoro-5,6-dihydro-5-methyl-6-oxo-4H-imidazo[1,5-a]benzodiazepine-3-carboxylate).

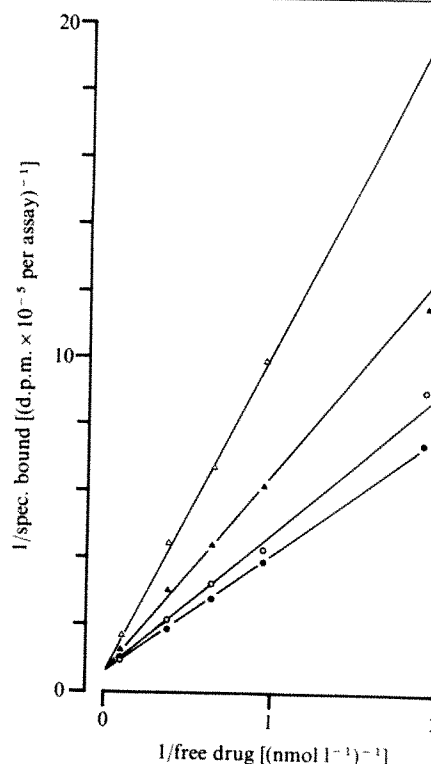


Fig. 2 Inhibition of specific binding of ³H-diazepam (0.5–10 nmol l⁻¹) in rat cerebral cortex homogenate by Ro 15-1788. ○, 0.5 nmol l⁻¹; ▲, 2 nmol l⁻¹; △, 5 nmol l⁻¹; ●, no Ro 15-1788.

augmented the dorsal root potential (index of GABA-mediated primary afferent depolarization) and depressed polysynaptic ventral root reflexes and γ -motoneurone activity. These effects were dose dependently reversed by i.v. injections of Ro 15-1788. In non-anaesthetized 'encéphale isolé' rats, i.v. injections of the short-acting imidazobenzodiazepine, midazolam⁷, produced a dose-dependent decrease in the spontaneous multi-unit activity in the substantia nigra pars compacta and in the nucleus raphé dorsalis; a 50% decrease was achieved between 326 and 750 μ g per kg. Ro 15-1788 at 3 mg per kg i.v. reversed the effect of the highest dose of midazolam on the firing rate of these supposed dopamine and serotonin neurones. In the same rats essentially the same antagonism between midazolam and Ro 15-1788 was found on hippocampal neurones.

Ro 15-1788 antagonized characteristic biochemical changes induced in the brain by benzodiazepines (H.M. *et al.*, in preparation). These drugs consistently attenuate the neuroleptic-induced increase of homovanillic acid (HVA) in the rat brain (an index of dopamine turnover) by enhancing a GABAergic inhibitory influence on dopamine neurones¹¹. Ro 15-1788 dose dependently reversed the reduction by diazepam of the chlorpromazine-induced increase of HVA. Benzodiazepines and sedative drugs of diverse chemical classes decrease the cyclic 3',5'-guanosine monophosphate (GMP) content of the cerebellum in rats^{12,13}. Ro 15-1788 by itself did not affect the cerebellar cyclic GMP level, but dose dependently prevented (when given before) and attenuated or abolished (when given afterwards) the cyclic GMP reduction induced by diazepam. The reduction of cerebellar cyclic GMP by barbiturates, ethanol, meprobamate, muscimol and neuroleptics was unaffected by Ro 15-1788.

The antagonistic activity of Ro 15-1788 is not only selective for active benzodiazepines as distinct from non-benzodiazepine depressants but is also selective for benzodiazepine effects mediated by the so-called central type of benzodiazepine receptor as distinct from effects produced through peripheral types of receptor or through nonspecific interactions with

biological systems. For example, the antischistosomal effect of Ro 11-3128 (refs 14, 15) was not inhibited by Ro 15-1788 *in vitro* and *in vivo* (H. R. Stohler, personal communication); the antischistosomal activity of some 7-nitro-substituted benzodiazepine derivatives is mediated by a low-affinity binding site in the epidermis of schistosomes with a different ligand specificity from that of the central type of benzodiazepine receptor¹⁶. Ro 15-1788 also failed to antagonize the depression by benzodiazepines in high concentrations (10^{-5} – 10^{-3} mol l⁻¹) of contractile responses of the guinea pig ileum to transmural electrical stimulation, an effect which is probably due to a nonspecific membrane stabilizing action.

These results characterize Ro 15-1788 as a representative of a novel class of drugs that act as specific antagonists of CNS effects of active benzodiazepines as well as of non-benzodiazepines with affinity for the central type of benzodiazepine receptors. The discovery of such antagonists is analogous to that of opiate antagonists and has considerable practical and scientific implications. They should be useful as antidotes for intoxications involving overdoses of benzodiazepines. As non-toxic agents, they should be useful for abolishing or reducing central effects of benzodiazepines given for surgical or diagnostic interventions whenever an immediate reversal of benzodiazepine effects is required. Previous attempts to counteract benzodiazepine effects by cholinesterase inhibitors^{17,18} or naloxone^{19,20} have yielded erratic results. Diagnostically, benzodiazepine antagonists should easily unmask the presence of active amounts of benzodiazepines in the CNS. Ro 15-1788 will be used to prevent the undesired central effects of Ro 11-3128 (3-methylclonazepam), which has proved effective as a one-shot treatment of schistosomiasis. Indeed, pilot studies in human volunteers have demonstrated a good tolerance and lack of intrinsic pharmacological effects of Ro 15-1788 as well as its ability to antagonize sedation induced by Ro 11-3128 (A. Darragh, personal communication). As research tools, Ro 15-1788 and other specific benzodiazepine antagonists will be extremely useful in defining the structural properties required for molecules to be recognized and bound by benzodiazepine receptors and to induce or prevent those conformational changes in the receptor protein that result in the typical enhancement of GABAergic transmission. Benzodiazepine antagonists should also be powerful tools in studying the possible existence of endogenous ligands for benzodiazepine receptors. Our screening for benzodiazepine antagonistic activity resulted in the identification of many compounds possessing either pure antagonistic or antagonistic properties combined with weak agonistic activity and revealed interesting and unexpected structure-activity relationships.

We thank Drs W. P. Burkard, M. Gerecke, H. H. Keller, E. Kyburz, J. P. Laurent and R. K. M. Müller for their contributions.

Received 7 November 1980; accepted 22 January 1981.

- Haefely, W. *et al.* in *Mechanism of Action of Benzodiazepines* (eds Costa, E. & Greengard, P.) 131–151 (Raven, New York, 1975).
- Costa, E., Guidotti, A., Mao, C. C. & Suria, A. *Life Sci.* **17**, 167–186 (1975).
- Braestrup, C. & Squires, R. F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3805–3809 (1977).
- Möhler, H. & Okada, T. *Science* **198**, 849–851 (1977).
- Möhler, H. *Trends Pharmac.* (in the press).
- Guidotti, A., Toffano, G. & Costa, E. *Nature* **275**, 553–555 (1978).
- Pieri, L. *et al.* *Drug Res.* (in the press).
- Boissier, J.-R. & Simon, P. *Thérapie* **15**, 1170–1174 (1960).
- Blanchard, J. C., Boireau, A., Garret, C. & Julou, L. *Life Sci.* **24**, 2417–2420 (1979).
- Polc, P., Möhler, H. & Haefely, W. *Naunyn-Schmiedeberg's Archs Pharmac.* **284**, 319–337 (1974).
- Keller, H. H., Schaffner, R. & Haefely, W. *Naunyn-Schmiedeberg's Archs Pharmac.* **294**, 1–7 (1976).
- Burkard, W. P., Pieri, L. & Haefely, W. *Adv. biochem. Psychopharmac.* **15**, 315–324 (1976).
- Biggio, G., Costa, E. & Guidotti, A. *Adv. biochem. Psychopharmac.* **15**, 325–335 (1976).
- Stohler, H. R. in *Current Chemotherapy* (eds Siegenthaler, W. & Lüthy, R.) 147–148 (American Society for Microbiology, Washington DC, 1978).
- Pax, R., Bennett, J. L. & Fetterer, R. *Naunyn-Schmiedeberg's Archs Pharmac.* **304**, 309–315 (1978).
- Bennett, J. L. *J. Parasit.* (in the press).
- Di Liberti, J. & O'Brien, M. L. *J. Pediatr.* **86**, 106–107 (1975).
- Ruppreht, J. *Anesthesiology* **53**, 180–181 (1980).
- Bell, E. F. *J. Pediatr.* **87**, 803–804 (1975).
- Christensen, K. N. & Hüttel, M. *Anesthesiology* **51**, 187 (1979).

Purified EGF receptor-kinase interacts specifically with antibodies to Rous sarcoma virus transforming protein

Michael Chinkers & Stanley Cohen

Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

Transformation by several RNA tumour viruses seems to be mediated by virally coded protein kinases which specifically phosphorylate tyrosine^{1–7}. A tyrosine-specific protein kinase also seems to be involved in the mitogenic action of epidermal growth factor (EGF)⁸. This EGF-stimulated kinase activity is closely associated with the EGF receptor, with which it copurifies during EGF-affinity chromatography⁹. Because both the virus- and EGF-stimulated tyrosine kinases may be involved in stimulation of cell growth, and because the viral kinases may be antigenically related to normal cell proteins^{5,10–12}, we examined the interaction of antibodies to viral tyrosine kinases with the affinity-purified EGF receptor-kinase preparation. We report here that the receptor-kinase specifically phosphorylates antibodies directed against the transforming protein kinase pp60^{src} of Rous sarcoma virus. However, none of these antibodies, including those which cross-react with the normal cellular homologue of pp60^{src} (pp60^{src}), precipitate the receptor-kinase. These results suggest that the EGF receptor-kinase is related to, but probably not identical with, pp60^{src}.

Antisera from rabbits bearing tumours induced by Rous sarcoma virus (TBR sera) not only precipitate pp60^{src} (ref. 13), but also have the unusual property of acting as a specific substrate for the protein kinase activity of pp60^{src}. In this reaction, pp60^{src} catalyses the phosphorylation of the IgG heavy chain^{14,15}, but does not phosphorylate normal rabbit IgG^{16,17}; TBR IgGs are not substrates for protein kinases in general¹¹. Using TBR sera obtained from several laboratories, we examined the ability of the EGF receptor-kinase to phosphorylate specifically the IgG heavy chain of TBR sera.

Affinity-purified EGF receptor-kinase, prepared from membranes of A431 cells⁹, was incubated with various antisera in the presence of EGF, Mn²⁺ and [γ -³²P]ATP and subjected to SDS-gel electrophoresis and autoradiography (Fig. 1). As previously reported, in the absence of added serum the receptor-kinase catalyses the phosphorylation of two bands with apparent molecular weights (MWs) of 150,000 and 170,000 (Fig. 1, lane a)—this doublet is thought to be the EGF receptor itself⁹. When TBR serum is included in the reaction mixture, a band comigrating with the IgG heavy chain becomes phosphorylated (Fig. 1, lanes b–f). This phosphorylation seems to be more pronounced when TBR sera that cross-react with the normal cellular homologue (pp60^{src}) of pp60^{src} are used (Fig. 1, lanes d–f). Normal rabbit sera are not phosphorylated or are phosphorylated only to a small extent (Fig. 1, lanes g, h). We tested 30 antisera to various proteins and viruses in this assay, and, with the exception of TBR sera, none was phosphorylated by the EGF receptor-kinase. These controls included antisera which precipitate the tyrosine kinases associated with Abelson murine leukaemia virus¹, feline sarcoma virus⁵ and polyoma virus¹⁸, as well as an antiserum to A431 cell membranes that precipitates the receptor-kinase (Fig. 1, lane i)¹⁹. Thus, the EGF receptor-kinase interacts specifically with antisera that precipitate pp60^{src}.

The observed phosphorylation of TBR IgG was not due to endogenous protein kinase activity in the TBR sera or to an EGF-unrelated cellular kinase in the purified receptor prep-

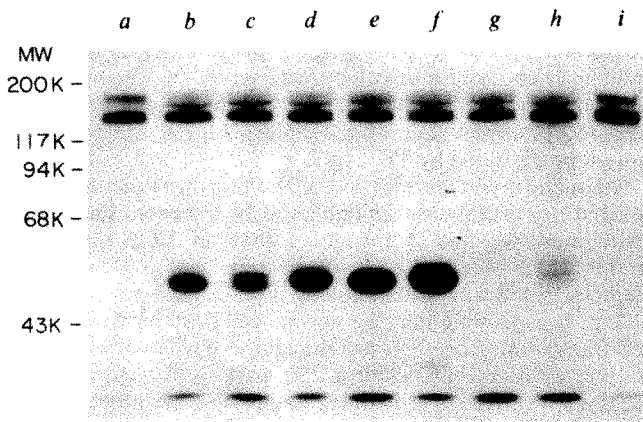


Fig. 1 Analysis by SDS-gel electrophoresis and autoradiography of EGF receptor-kinase-catalysed phosphorylation of rabbit anti-tumour sera. Antisera from rabbits bearing tumours induced by Rous sarcoma virus (TBR sera)¹³ were able to precipitate and act as a substrate for the transforming protein kinase coded for by Rous sarcoma virus (pp60^{src})^{14,15}. The antisera shown in lanes *d-f* can also precipitate, and be phosphorylated by, the homologous cellular pp60^{src} (ref. 11). Rabbit antiserum raised against A431 human epidermal carcinoma cell membranes was prepared and purified as described elsewhere¹⁹; this antiserum precipitates the purified EGF receptor with its associated tyrosine-specific⁸ protein kinase activity (EGF receptor-kinase). The EGF receptor-kinase was purified from Triton/glycerol-solubilized A431 cell membranes by adsorption to EGF-AffiGel and elution with ethanolamine as described elsewhere⁹. The receptor-kinase preparation used in these experiments contained ~90 µg ml⁻¹ of protein in 1% Triton X-100, 10% glycerol, 5 mM ethanolamine, pH 7. 15 µl of receptor-kinase was activated⁹ by incubation for 10 min at 25 °C with 15 µl EGF (100 ng, in 0.05% bovine serum albumin (BSA)) and 15 µl of 4 mM MnCl₂ in 60 mM HEPES, pH 7.4 (HEPES/Mn²⁺). The mixture was then chilled on ice for 10 min and incubated with 3 µl of the indicated antiserum for 30 min at 0 °C. 15 µl [γ -³²P]ATP (2 µCi, 60 µM) was then added, and after 10 min at 0 °C the phosphorylation reaction was terminated by addition of SDS sample buffer to the final concentrations described by Laemmli²¹. Samples were heated for 5 min in a boiling water bath and electrophoresed through a 3% polyacrylamide stacking gel and a 7.5% separating gel using Laemmli's buffer system²¹. ³²P-labelled proteins were visualized by autoradiography on DuPont Cronex film, while molecular weight markers (BioRad) and the IgG heavy chain were located by Coomassie blue staining. Markers were myosin, MW 200,000 (200 K); β -galactosidase, MW 117 K; phosphorylase *b*, MW 94 K; BSA, MW 68 K, and ovalbumin, MW 43 K. Serum additions were as follows: *a*, saline control; *b-f*, TBR sera; *g*, *h*, normal rabbit sera; *i*, rabbit anti-A431 cell membrane IgG (25 µg).

ation (Fig. 2) because no phosphorylation of IgG was observed in any of the TBR sera in the absence of added receptor-kinase (Fig. 2, lane *a*), and phosphorylation of the IgG heavy chain was stimulated by EGF (Fig. 2, lanes *b*, *c*).

Surprisingly, despite the specific interaction of the EGF receptor-kinase with TBR IgG, none of the TBR sera examined was able to precipitate the receptor-kinase, including antisera known to precipitate pp60^{src} from uninfected cells (Fig. 3). The receptor doublet and the IgG heavy chain were labelled with ³²P by incubating receptor-kinase with TBR serum in the presence of EGF, [γ -³²P]ATP and Mn²⁺ (Fig. 3, lane *a*). This prelabelled mixture was then subjected to immunoprecipitation with protein A-Sepharose, and the supernatant (Fig. 3, lane *b*) and washed pellet (Fig. 3, lane *c*) were examined by SDS-gel electrophoresis and autoradiography. The phosphorylated IgG is precipitated but the receptor remains in the supernatant. However, as it is possible that the receptor and kinase are dissociable, this experiment was not adequate to determine

whether the receptor-associated kinase activity had remained in the supernatant or had been precipitated with the TBR IgG. Therefore, unlabelled EGF receptor-kinase, pre-activated with EGF, was incubated with unlabelled TBR serum, immunoprecipitated with protein A-Sepharose, and the supernatant and washed pellet were assayed for protein kinase activity in the presence of [γ -³²P]ATP and Mn²⁺. All phosphorylating activity was found in the supernatant, where the receptor doublet became phosphorylated (Fig. 3, lane *d*), whereas the IgG-containing precipitate contained no detectable kinase activity (Fig. 3, lane *e*). Thus, neither the EGF receptor nor its associated protein kinase activity is precipitated by TBR serum. The antisera to other viral tyrosine kinases (see above) also failed to precipitate the EGF receptor (not shown). As a positive control, it was confirmed that this TBR serum could precipitate IgG kinase activity from a crude detergent extract of A431 cells (Fig. 3, lane *g*). Thus, although A431 cells seem to contain pp60^{src} that precipitates with and phosphorylates TBR serum, this kinase is not detected in the purified receptor preparation. Furthermore, in these conditions both the receptor and its associated kinase activity could be immunoprecipitated using antiserum directed against the A431 cell membranes from which the receptor-kinase was purified (Fig. 3, lane *f*). In this experiment, unlabelled receptor-kinase, precipitated with the anti-membrane serum, became labelled in the immunoprecipitate on addition of [γ -³²P]ATP and Mn²⁺; the antibody was not labelled.

Because phosphorylation of IgG catalysed by pp60^{src} and phosphorylation of the 150,000–170,000 MW EGF receptor doublet by the receptor-kinase are known to occur on tyrosine residues^{2,8}, it was interesting to determine whether the IgG phosphorylation catalysed by the receptor-kinase was also specific for tyrosine (Fig. 4). TBR IgG was phosphorylated in the presence of the receptor-kinase and [γ -³²P]ATP, and precipitated with protein A-Sepharose; the precipitated material was eluted from the washed beads by heating in SDS, and SDS-gel electrophoresis and autoradiography of an aliquot of this eluate were used to confirm that the only major phosphorylated component was the IgG heavy chain (Fig. 4*a*). The remainder of the eluted material was subjected to partial acid hydrolysis and paper electrophoresis at pH 3.5 in the presence of internal phosphoamino acid standards⁸. With or without EGF,

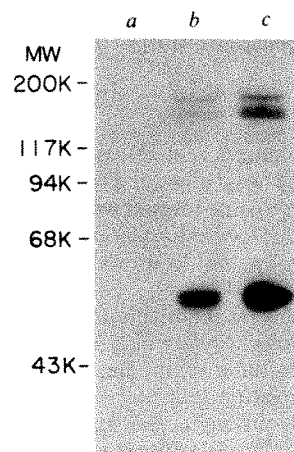


Fig. 2 Effect of EGF on receptor-kinase-catalysed phosphorylation of TBR serum. Phosphorylation reactions, electrophoresis and autoradiography were carried out as described in the legend to Fig. 1, except that buffer blanks were substituted for EGF or for EGF receptor-kinase as indicated. *a*, Minus receptor-kinase, +EGF, +TBR serum; *b*, plus receptor-kinase, -EGF, +TBR serum; *c*, plus receptor-kinase, +EGF, +TBR serum. Similar results (not shown) were obtained with the other TBR sera described in Fig. 1 legend.

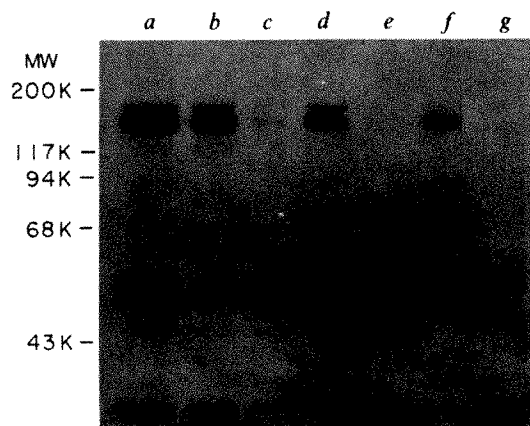


Fig. 3 SDS-gel electrophoresis and autoradiography showing non-precipitation of EGF receptor-kinase by TBR serum. *a-c*, Mixtures of EGF receptor-kinase and TBR serum (known to cross-react with pp60^{src}) were phosphorylated as described in Fig. 1 legend, except that the final reaction volume was increased to 83 μ l. *a*, This phosphorylated mixture was subjected to SDS-gel electrophoresis without further treatment; *b, c*, IgG in a replicate phosphorylated mixture was adsorbed to 40 μ l of packed protein A-Sepharose in 20 mM HEPES, pH 7.4 by incubation for 30 min at 0°C with resuspension every 5 min. The sample was then spun for 2 min at 8,000g in an Eppendorf minifuge, and 60 μ l of the supernatant prepared for electrophoresis (lane *b*). The protein A-Sepharose beads were washed five times with 1 ml of 20 mM HEPES, pH 7.4. The bound IgG was eluted by heating for 5 min at 100°C in 150 μ l of SDS sample buffer, and, after removal of the protein A-Sepharose beads by centrifugation, was subjected to SDS-gel electrophoresis (lane *c*). *d, e*, Unlabelled receptor-kinase was activated with EGF and incubated with TBR serum as in *a-c*. Following incubation, IgG was precipitated with protein A-Sepharose as described above for lanes *b* and *c*; 60 μ l of the supernatant was incubated for 10 min at 0°C with 20 μ l each HEPES/Mn²⁺ and [γ -³²P]ATP (2.0 μ Ci, 60 μ M) and processed for electrophoresis (lane *d*). The protein A-Sepharose precipitate was washed five times with 1 ml of 20 mM HEPES, pH 7.4, and incubated for 10 min at 0°C with 20 μ l each HEPES/Mn²⁺ and [γ -³²P]ATP. IgG bound to the protein A-Sepharose was eluted as for *c* and electrophoresed (lane *e*). *f, g*, Positive controls for immunoprecipitation: *f*, as lane *e*, except that rabbit IgG raised against A431 cell membranes was used in place of TBR serum; *g*, a crude 1% Triton/10% glycerol extract of unlabelled A431 cells was incubated with cross-reactive TBR serum and immunoprecipitated. Precipitation with protein A-Sepharose and phosphorylation of the precipitate with [γ -³²P]ATP was as for samples *e* and *f*.

the major phosphoamino acid released by hydrolysis of the phosphorylated IgG was identified as phosphotyrosine; the amount of phosphotyrosine formed was greater in the presence of the hormone. Thus, phosphorylation of TBR IgG by either pp60^{src} (ref. 2) or by the EGF receptor-kinase occurs on tyrosine.

The previous findings that both pp60^{src} (refs 2, 4) and the EGF receptor-kinase⁸ phosphorylate only tyrosine residues suggested some possible similarity in the active sites of the two enzymes. Our observation that both kinases specifically phosphorylate TBR IgG, although other IgGs contain tyrosine residues, suggests further similarity. However, there are several interpretations of the degree of this similarity: (1) TBR IgG, although not substrates for cyclic AMP-dependent protein kinases¹¹, may contain sequences that make them efficient substrates for a variety of related tyrosine-specific protein kinases. In this regard, it is interesting that the tyrosine kinase activities associated with the transforming proteins of feline sarcoma virus, Abelson murine leukaemia virus, polyoma virus, and the avian sarcoma viruses Y73 and PRC II are able specifically to phosphorylate TBR IgG (K. Beemon and T. Hunter, personal communication). (2) The EGF receptor-kinase may be

sufficiently antigenically related to pp60^{src} specifically to bind and phosphorylate anti-pp60^{src} IgG, but insufficiently related to bind tightly enough for immunoprecipitation. (3) The receptor-kinase may be pp60^{src}, but the interaction of pp60^{src} with the EGF receptor may inhibit the ability of the kinase to be immunoprecipitated by TBR IgG.

Although further purification of the receptor-kinase will be required to distinguish unambiguously between the above possibilities, we find it intriguing that an EGF-stimulated enzyme which seems to be involved in normal growth control interacts specifically with antibodies to a viral transforming protein. It is possible that the normal cell proteins from which viral transforming proteins are thought to have evolved²⁰ may, like the EGF receptor-kinase, be part of the normal cell machinery involved in growth regulation. Furthermore, as EGF is not tumorigenic, we believe that study of the means by which different tyrosine kinases stimulate growth may elucidate the cellular controls that distinguish malignant from non-malignant growth.

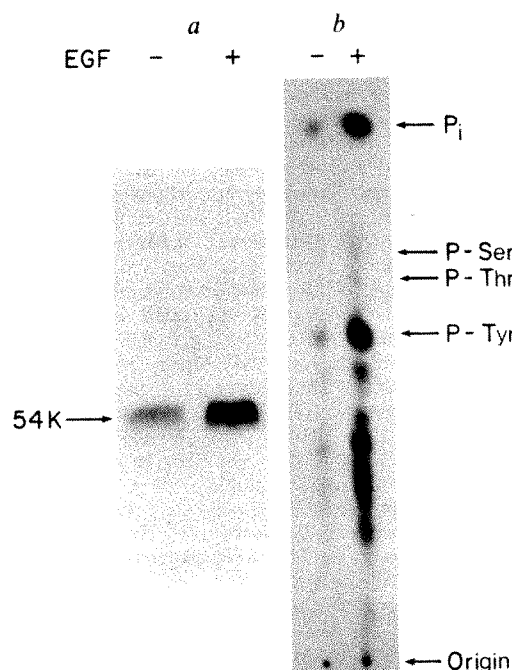


Fig. 4 Identification of phosphotyrosine in TBR IgG phosphorylated by the EGF receptor-kinase in the absence/presence of EGF. 40 μ l of the receptor-kinase preparation was mixed with 40 μ l of HEPES/Mn²⁺ and 40 μ l of either 0.05% BSA without EGF, or 0.05% BSA containing 270 ng EGF, and was incubated for 10 min at 25°C. The mixture was chilled on ice for 10 min and incubated for 1.5 h at 0°C with 6 μ l of TBR serum. 4.5 μ l of [γ -³²P]ATP (35 μ Ci, 436 μ M) was then added, and incubation on ice was continued for 10 min. 40 μ l of protein A-Sepharose in 20 mM HEPES, pH 7.4 was then added, and the suspension incubated for 30 min on ice, with resuspension every 5 min. The protein A-Sepharose was washed five times with 1 ml of 20 mM HEPES, pH 7.4 by spinning for 2 min at 8,000g in an Eppendorf minifuge, and the bound IgG was eluted by adding 600 μ l of 2% SDS in 20 mM HEPES, pH 7.4 and heating for 5 min in a boiling water bath. Aliquots (50 μ l) of each sample were examined by SDS-gel electrophoresis and autoradiography (*a*). The remainders of the samples were transferred to glass centrifuge tubes on ice and precipitated with 20% trichloroacetic acid (TCA) in the presence of 100 μ g BSA and 10 mM sodium pyrophosphate. Precipitates were washed twice with 20% TCA/10 mM sodium pyrophosphate and once each with ethanol and ether. Samples were air dried, hydrolysed under vacuum in 6 M HCl for 2.5 h at 100°C, and processed for paper electrophoresis at pH 3.5, as described elsewhere⁸. Results of paper electrophoresis and autoradiography are shown in *b*. Internal phosphoamino acid standards⁸ were located by staining with ninhydrin.

We thank Drs J. M. Bishop, D. Baltimore, T. Hunter, M. Barbacid and R. Erikson for their gifts of antisera, and Martha Reich and Katherine Chen for technical assistance. This work was supported by USPHS grant HD-00700. S.C. is an American Cancer Society Research Professor.

Received 7 October 1980; accepted 26 January 1981.

- Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826-831 (1980).
- Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
- Sefton, B. M., Hunter, T., Beemon, K. & Eckhart, W. *Cell* **20**, 807-816 (1980).
- Collet, M. S., Purchio, A. F. & Erikson, R. L. *Nature* **285**, 167-169 (1980).
- Barbacid, M., Beemon, K. & Devare, S. G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5158-5162 (1980).
- Reynolds, F. H. Jr, Van de Ven, W. J. M. & Stephenson, J. R. *J. biol. Chem.* **255**, 11040-11047 (1980).
- Kawai, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6199-6203 (1980).
- Ushiro, H. & Cohen, S. *J. biol. Chem.* **255**, 8363-8365 (1980).
- Cohen, S., Carpenter, G. & King, L. Jr *J. biol. Chem.* **255**, 4834-4842 (1980).
- Collett, M. S., Brugge, J. S. & Erikson, R. L. *Cell* **15**, 1363-1369 (1978).
- Oppermann, H., Levinson, A. D., Varmus, H. E., Levintow, L. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1804-1808 (1979).
- Witte, O. N., Rosenberg, N. E. & Baltimore, D. *Nature* **281**, 396-398 (1979).
- Brugge, J. S. & Erikson, R. L. *Nature* **269**, 346-348 (1977).
- Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021-2024 (1978).
- Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. M. *Cell* **15**, 561-572 (1978).
- Maness, P. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 5028-5032 (1979).
- Erikson, R. L., Collett, M. S., Erikson, E. & Purchio, A. F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6260-6264 (1979).
- Eckhart, W., Hutchinson, M. A. & Hunter, T. *Cell* **18**, 925-933 (1979).
- Haigler, H. T. & Carpenter, G. *Biochim. biophys. Acta* **598**, 314-325 (1980).
- Hanafusa, H. in *Comprehensive Virology* Vol. 10 (eds Fraenkel-Conrat, H. & Wagner, R. P.) 401-483 (Plenum, New York, 1977).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Anti-pp60^{src} antibodies are substrates for EGF-stimulated protein kinase

Jeffrey E. Kudlow, Janice E. Buss & Gordon N. Gill

Department of Medicine, Division of Endocrinology, University of California, San Diego, School of Medicine, La Jolla, California 92093, USA

Epidermal growth factor (EGF) stimulates phosphorylation of its own receptor at a tyrosine residue¹. Similarly, the viral gene product pp60^{src}, which is responsible for cellular transformation by avian sarcoma virus (ASV), phosphorylates itself and immunoglobulin directed against pp60^{src} at tyrosine residues^{2,3}. This unusual site of phosphorylation catalysed by two membrane-associated protein kinases⁴⁻⁷ involved in growth control prompted us to study the immunological relatedness of the EGF-stimulated protein kinase and the pp60^{src}. Using anti-pp60^{src} antisera, we attempted to immunoprecipitate the EGF-stimulated protein kinase solubilized from plasma membranes. We report here that neither the EGF-stimulated kinase nor the EGF receptor were immunoprecipitable by anti-pp60^{src} sera. However, anti-pp60^{src} IgG served as a specific substrate for the EGF-stimulated kinase, suggesting a close similarity between the EGF-stimulated kinase and pp60^{src}.

We used the plasma membranes of A431 human carcinoma cells because of their unusually high density of EGF receptors⁸. Detergent-solubilized plasma membranes prepared from these cells retain EGF-stimulated phosphorylation of the EGF receptor protein⁹ of molecular weight (MW) ~170,000. Treatment of the solubilized membranes with antisera raised in ASV tumour-bearing rabbits¹⁰ was followed by removal of the immune complexes with protein A containing *Staphylococcus aureus* (Pansorbin, Calbiochem). Phosphorylation of immune, but not control, IgG heavy chains could be demonstrated in the resultant precipitate. Normal or non-ASV-transformed vertebrate cells contain *src* gene sequences and a cell protein

(pp60^{src}) that is closely related to the viral pp60^{src} (refs 11-14). Our ability to precipitate specifically a rabbit IgG heavy chain kinase with anti-pp60^{src} sera suggested that the A431 plasma membrane contained pp60^{src}.

Using this immunoprecipitation approach, we determined the effect of removal of pp60^{src} activity on EGF-stimulated phosphorylation of the EGF receptor. The membranes were solubilized using a procedure that optimized the recovery of pp60^{src} activity in the immunoprecipitates and the solubilized A431 membranes were then titrated with antiserum to determine the amount required to remove the maximum quantity of pp60^{src} activity made available by membrane solubilization. As determined by this titration, doubling of the amount of antiserum did not increase the amount of pp60^{src} activity precipitated. However, another incubation with additional antiserum in the immune adsorbed supernatant detected some residual pp60^{src} activity (Fig. 1, lane 8), which never exceeded 25% of the total immunoprecipitable activity. A third immune precipitation failed to detect any residual pp60^{src} activity. When immunoprecipitates were formed from solubilized plasma membranes

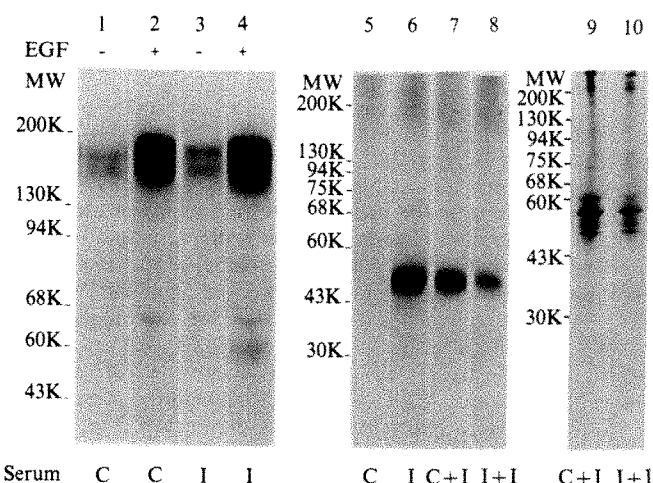


Fig. 1 Effect of immunoprecipitation of pp60^{src} activity on EGF-stimulated phosphorylation of the EGF receptor in solubilized A431 cell membranes. Purified A431 plasma membranes¹⁷ were solubilized in buffer containing 1% Triton X-100, 10% glycerol, and 20 mM HEPES, pH 7.4 (ref. 9). Control (lanes 1, 2) and anti-pp60^{src} (lanes 3, 4) sera at a final dilution of 1/20 were incubated with 400 µg ml⁻¹ of solubilized A431 membrane for 3 h in 20 mM HEPES, pH 7.4, 1 mM MnCl₂, and 0.1 mg ml⁻¹ bovine serum albumin (BSA). Control (lane 5) and immune (lane 6) IgG were removed from the mixture by precipitation with Pansorbin for quantitation of pp60^{src} activity and the remaining supernatants divided into two aliquots: one was treated with (lanes 2, 4) or without (lanes 1, 3) 0.3 µM EGF (ref. 9) followed by the addition of 15 µM [γ -³²P]ATP (1.25 µCi per 60 µl assay). After incubation for 2 min at 0°C, the reaction was stopped by the addition of Laemmli sample buffer¹⁸ and boiling for 3 min, followed by electrophoresis on 7.5% polyacrylamide gels. Phosphorylated proteins were visualized by autoradiography. To quantify residual pp60^{src} present in control and immune supernatants, second aliquots of supernatant were again treated with anti-pp60^{src} serum and Pansorbin. These second immune precipitates derived from control (lane 7) or immune (lane 8) serum-treated supernatants were assayed for pp60^{src} activity. The Pansorbin-bound immune precipitates were washed three times by centrifugation in NP40 buffer¹⁹, once in phosphate-buffered saline, and resuspended in 80 µl of a buffer containing 20 mM sodium phosphate (pH 7.2), 5 mM MgCl₂ and 15 nM [γ -³²P]ATP (3,000 Ci mmol⁻¹, 2 µCi per assay). After incubation for 20 min at 30°C, the reaction was stopped by the addition of RIPA buffer²⁰, and phosphorylated proteins were separated on 10% polyacrylamide gels. The peptides present in the immunoprecipitates were analysed (lanes 9, 10). A431 cells were metabolically labelled with ³⁵S-methionine¹⁴, the plasma membranes isolated and solubilized and immunoprecipitations performed. The Pansorbin-bound immunoprecipitates from solubilized membranes that had been pre-adsorbed with control (lane 9) or immune (lane 10) serum were washed extensively before separation on 10% polyacrylamide gels. Molecular weight standards were: myosin, MW 200,000 (200 K); β -galactosidase, 130 K; phosphorylase b, 94 K; transferrin, 75 K; BSA, 68 K; catalase, 60 K; ovalbumin 43 K, and carbonic anhydrase, 30 K. C designates treatment with control serum, I with anti-pp60^{src} serum and C+I or I+I, sequential treatment with these sera.

Table 1 EGF-stimulated phosphorylation of tyrosine residues in the EGF receptor and in anti-pp60^{src} IgG

	³² P incorporation into tyrosine residues (c.p.m.) EGF receptor (MW ~ 170,000)	IgG heavy chain (MW = 55,000)
-EGF	156	147
+EGF	349	513

Phosphorylation reaction mixtures containing solubilized A431 membranes and anti-pp60^{src} serum with or without 0.3 μ M EGF were incubated as described in the legends to Figs 1 and 2. After electrophoresis, the gels were dried and the phosphorylated EGF receptor and IgG heavy chain located by autoradiography using appropriate alignment markers. The regions of the dried gel containing the phosphoproteins were cut out and placed in dialysis bags after removal of the paper backing. Buffer containing freshly prepared 10 mM NH_4HCO_3 , 0.1% SDS and 5% 2-mercaptoethanol was placed in the dialysis bag to swell the gel. The dialysis bags were placed in the same buffer between two platinum electrodes and a 5–10 V cm^{-1} field applied for 8 h; the field was reversed for 15 s before removal of the contents. This solution was filtered through glass wool and 50 μ g of bovine γ -globulin carrier added before trichloroacetic acid (TCA) precipitation. The protein precipitate was washed in ethanol and ether and hydrolysed for 2 h in 6 M HCl at 110 °C. Phosphoamino acids were separated by electrophoresis on cellulose thin layer plates with the buffer (pH 3.5) described by Hunter and Sefton⁷. ³²P-labelled amino acids were located by autoradiography and identified by alignment with the corresponding ninhydrin-stained phosphoamino acid standards. The area of the thin layer plate corresponding to phosphotyrosine was scraped and its ³²P content determined by scintillation counting.

prepared from ³⁵S-methionine-labelled A431 cells, several peptides were precipitated nonspecifically (data not shown). However, if immunoprecipitates were formed from solubilized membranes pre-adsorbed with control serum, a peptide of MW 60,000 was evident. Pre-adsorption with anti-pp60^{src} serum markedly and specifically diminished the amount of the 60,000-MW peptide in the immunoprecipitate, approximately in proportion to the removal of anti-pp60^{src} IgG kinase activity (Fig. 1, lanes 9, 10). Despite specific removal of >75% of the pp60^{src} activity and the associated 60,000-MW peptide from the solubilized membrane, EGF-stimulated phosphorylation of the ~170,000-MW receptor protein was unaffected (Fig. 1, lanes 1–4). Furthermore, analysis of the products of phosphorylation of the immunoprecipitate never revealed the ~170,000-MW receptor (Fig. 1, lanes 6–8) nor could EGF receptor phosphorylated before immunoprecipitation be detected in the immunoprecipitates (See Fig. 3b). Thus, neither the EGF receptor nor the EGF-stimulated protein kinase was removed by immunoprecipitation. These data suggested that either the pp60^{src} was immunologically unrelated to the EGF-stimulated kinase or that it was unavailable for immune adsorption.

In view of these experiments, we were surprised with the results obtained when solubilized A431 membranes were incubated with antisera which were not subsequently removed with Pansorbin: in these conditions, both the receptor and the heavy chains of anti-pp60^{src} IgG, but not control IgG, were phosphorylated in an EGF-stimulated manner (Fig. 2, upper panel). Antisera raised in several ASV tumour-bearing rabbits were phosphorylated by the solubilized membranes to varying extents. Those antisera which were good substrates for the solubilized EGF-stimulated kinase were also good substrates in the immunoprecipitation reaction designed to detect pp60^{src} activity (Fig. 2, lower panel). Because these antisera were selected for their ability to be phosphorylated by pp60^{src}, they did not always have parallel ability to act as substrate for pp60^{src}. Various non-immune sera failed to serve as phosphorylation substrates for either kinase. The parallel ability of the immune antisera to act as substrates for pp60^{src} and the EGF-stimulated kinase suggested that very similar determinants are required by the pp60^{src} and the EGF-stimulated kinase to allow phosphorylation of the anti-pp60^{src} sera.

The phosphoamino acids in the receptor and IgG heavy chain were analysed in phosphoproteins purified by polyacrylamide gel electrophoresis. Both the receptor and IgG were phosphorylated exclusively at tyrosine residues, and phosphate

incorporation into tyrosine in both proteins was stimulated threefold by EGF (Table 1).

The EGF stimulation of IgG phosphorylation was not due to the release of more pp60^{src} into an immunoreactive fraction because EGF failed to increase significantly the yield of pp60^{src} activity in the immune precipitate (data not shown). Furthermore, removal of a significant fraction of pp60^{src} did not alter the ability of the EGF-stimulated kinase to phosphorylate immune IgG (Fig. 3a). The inability to alter the receptor or immune IgG phosphorylation by prior immunoprecipitation of pp60^{src} suggested that these proteins were not phosphorylated by the fraction of pp60^{src} accessible to immunoprecipitation. The failure to immunoprecipitate the EGF receptor or its EGF-stimulated kinase, which is tightly associated with the EGF receptor, was not due to any impediment to binding of the Fc fragment of IgG to the protein A on the bacteria. Pansorbin could quantitatively remove antibodies which had been interacted with the EGF-stimulated kinase and phosphorylated. However, such immune precipitation did not precipitate the phosphorylated EGF receptor (Fig. 3b).

The apparent low affinity of the anti-pp60^{src} IgG for the EGF-stimulated kinase, which allowed phosphorylation but not immune precipitation, has two possible explanations. (1) The EGF-stimulated kinase could be distinct from the pp60^{src} yet similar enough to allow immune-specific recognition of antisera leading to IgG heavy chain phosphorylation. A closer homology between pp60^{src} kinase and EGF-stimulated kinase cannot be excluded. The tight association between the kinase and receptor could occlude some immune determinants, allowing the anti-pp60^{src} IgG to interact strongly enough for phosphorylation, but too weakly for immune precipitation. (2) It is possible that the phosphorylation site in the variable region of the heavy chain¹⁵ of anti-pp60^{src} contains a sequence that is preferred by tyrosine-specific kinases.

In either case, a close similarity between pp60^{src} and the EGF receptor kinase has been established. Both proteins are themselves phosphorylated at tyrosine residues and both are able to

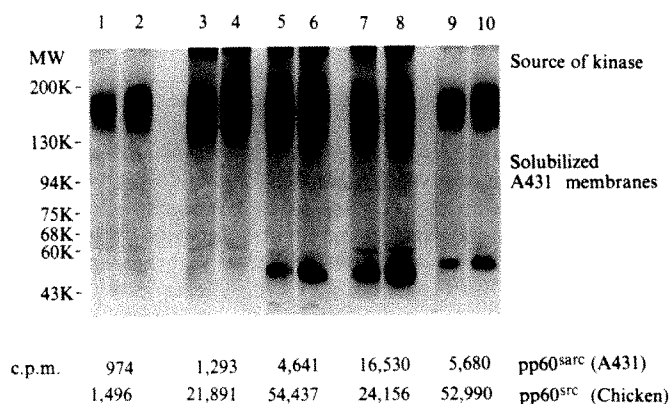


Fig. 2 EGF-stimulated phosphorylation of anti-pp60^{src} sera. Various antisera to viral pp60^{src} were raised in newborn rabbits bearing ASV-induced tumours¹⁰. Sera from newborn rabbits that did not develop tumours contained no detectable anti-pp60^{src} activity—these and several adult rabbit sera served as controls. Control sera (lanes 1, 2) and immune sera (lanes 3–10) were incubated with solubilized A431 membranes for 30 min at 0 °C as described in Fig. 1 legend, without the addition of Pansorbin. Reaction mixtures were treated with (lanes 2, 4, 6, 8, 10), or without (lanes 1, 3, 5, 7, 9), 0.3 μ M EGF followed by the addition of 15 μ M [γ -³²P]ATP for 2 min at 0 °C. Electrophoresis and autoradiography were performed as described in Fig. 1 legend. Numbers below the autoradiogram show results for the same antisera when tested for ability to precipitate, and be phosphorylated by, cellular pp60^{src} from solubilized A431 membranes and by partially purified pp60^{src} (ref. 15) from ASV-induced chicken tumours. Immunoprecipitations and phosphorylation reactions were carried out as in Fig. 1 legend except that the reactions were terminated by applying the reaction mixtures to Whatman 3M filter paper discs which were then placed in 10% TCA. The filters were washed three times in 5% TCA, then boiled in 5% TCA for 10 min; TCA and water were then removed by washing the filters in ethanol, followed by ether. ³²P incorporated into IgG heavy chains was measured by scintillation counting.

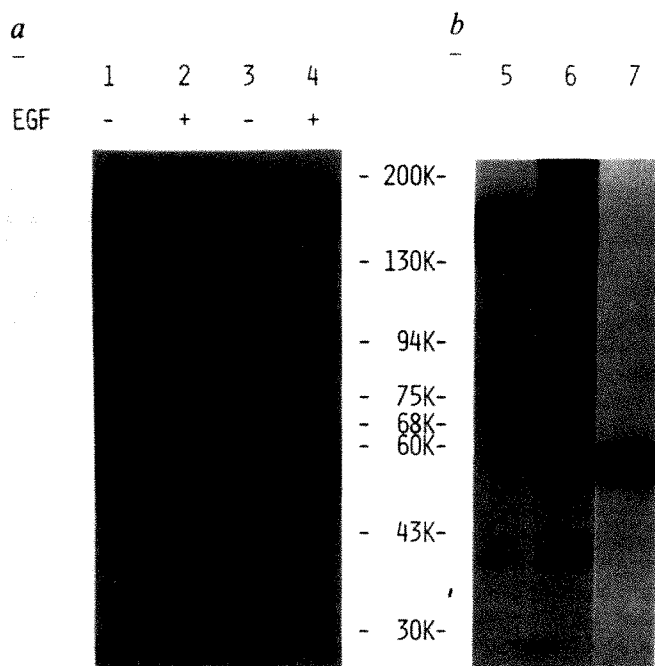


Fig. 3 *a*, Lack of effect of removal of pp60^{src} on EGF-stimulated phosphorylation of anti-pp60^{src} serum. pp60^{src} was removed from solubilized A431 membranes by immunoprecipitation with anti-pp60^{src} serum. The same quantity of antiserum was re-added to control (lanes 1, 2) and immune (lanes 3, 4) serum-treated supernatants. Mixtures were incubated for 30 min at 0°C followed by addition of EGF (lanes 2, 4) or buffer (lanes 1, 3). Phosphorylation reactions, electrophoresis and autoradiography were performed as described in Fig. 1 legend. *b*, Failure of anti-pp60^{src} serum to precipitate the phosphorylated EGF receptor. Immune serum and EGF were added to solubilized A431 cell membranes as described in Fig. 2 legend. [γ -³²P]ATP (15 μ M) was added for 5 min at 0°C and the reaction terminated by addition of EDTA to a final concentration of 5 mM. Immunoprecipitation was carried out using Pansorbin. Samples were analysed by electrophoresis and autoradiography. Lane 5, supernatant remaining after immunoprecipitation; lane 6, complete phosphorylation mixture; lane 7, immunoprecipitate.

phosphorylate, with parallel avidity, antisera directed against pp60^{src}, also at tyrosine residues. Although identity of the EGF-stimulated kinase with the pp60^{src} has not been established, these two kinases belong to a new family of tyrosine kinases involved in the continuum from growth control to transformation. It has been suggested that ASV transformation is an exaggerated growth response resulting from a 30- to 50-fold increase in the amount of pp60^{src} in the transformed cells¹². Although EGF stimulation does not result in transformation, sarcoma growth factor, described by DeLarco and Todaro¹⁶, does induce a transformation-like state in EGF responsive cells, apparently through binding to the EGF receptor. It is interesting to speculate that sarcoma growth factor acts as a super-agonist on the EGF receptor, increasing the tyrosine phosphorylation activity of the EGF-stimulated kinase to levels seen in ASV transformation.

We thank Drs Tony Hunter and Bartholemew Sefton for helpful advice and gifts of ASV, anti-pp60^{src} serum and phosphoamino acid standards. This work was supported by American Cancer Society research grant BC-209 and NIH research grant AM13149 from the National Institute of Arthritis, Metabolism and Digestive Diseases. J.E.K. is a Centennial Fellow of the MRC of Canada.

Received 7 October 1980; accepted 26 January 1981.

1. Ushiro, H. & Cohen, S. *J. biol. Chem.* **255**, 8363-8365 (1980).
2. Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
3. Collett, M. S., Purchio, A. F. & Erikson, R. L. *Nature* **285**, 167-169 (1980).
4. Brugge, J. S., Steinbaugh, P. J. & Erikson, R. L. *Virology* **91**, 130-140 (1978).

5. Rohrschneider, L. R. *Cell* **16**, 11-24 (1979).
6. Willingham, M. C., Jay, G. & Pastan, I. *Cell* **18**, 125-134 (1981).
7. Courtneidge, S. A., Levinson, A. D. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3783-3787 (1980).
8. Fabricant, R. N., DeLarco, J. E. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 565-569 (1977).
9. Cohen, S., Carpenter, G. & King, L. Jr *J. biol. Chem.* **255**, 4834-4842 (1980).
10. Brugge, J. S. & Erikson, R. L. *Nature* **269**, 346-348 (1977).
11. Collett, M. S., Brugge, J. S. & Erikson, R. L. *Cell* **15**, 1363-1370 (1978).
12. Karsess, R. E., Hayward, W. S. & Hanafusa, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3154-3158 (1979).
13. Oppermann, H., Levinson, A. D., Varmus, H. E., Levintow, L. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1804-1808 (1979).
14. Sefton, B. M., Hunter, T. & Beemon, K. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2059-2063 (1980).
15. Maness, P. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 5028-5032 (1979).
16. DeLarco, J. E. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4001-4005 (1978).
17. Thom, D., Powell, A. J., Lloyd, C. W. & Rees, D. A. *Biochem. J.* **168**, 187-194 (1977).
18. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
19. Sefton, B. M., H.
20. Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021-2024 (1978).

Human HLA gene segment isolated by hybridization with mouse H-2 cDNA probes

Bertrand R. Jordan*, François Bregegere† & Philippe Kourilsky†

* Centre d'Immunologie INSERM-CNRS de Marseille-Luminy, Case 906, 13288 Marseille Cédex 9, France

† Unité de Biologie Moléculaire du Gène, ER CNRS 201 and SCN INSERM 20, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cédex 15, France

The products of the highly polymorphic genes known as the major histocompatibility complex (MHC) have been shown to play a major part in the control of several aspects of the immune response and of susceptibility to certain diseases^{1,2}. The major transplantation antigens are membrane proteins composed of two noncovalently associated polypeptide chains: a light, non-polymorphic chain of molecular weight 12,000 (MW), called β_2 -microglobulin, and a polymorphic, glycosylated heavy chain of ~45,000 MW. The heavy chains in man and mouse are encoded in at least three loci of the MHC named HLA-A, -B and -C, and H-2D, K and L respectively. The available amino acid sequence data indicate extensive homology between human and mouse heavy chains^{3,4}. We therefore used a recently isolated cloned mouse H-2 cDNA probe⁵ to screen a human gene library. We have now characterized one of the recombinant phages obtained, λ HLA-12, and shown that it contains an authentic HLA sequence with evidence of a second one close by.

Major efforts have been directed towards the molecular cloning of HLA and H-2 gene sequences. The isolation of cloned cDNA copies has proved difficult because of the low abundance of the corresponding mRNAs *in vivo*. Nevertheless, cDNA clones carrying HLA⁶ and H-2 (ref. 5) sequences have been obtained. A recombinant plasmid, pH-2^{d-3}, recently isolated by Bregegere *et al.*⁷, carries a cDNA insert of 1,050 base pairs, which corresponds to the 3'-untranslated region (500 base pairs) and 550 base pairs of translated sequence of a mRNA coding for a H-2 heavy chain of the *d* haplotype. H-2 and HLA heavy chains consist of three globular domains followed by a transmembrane sequence which anchors the molecule at the cell surface, and a short intracellular carboxy-terminal fragment. pH-2^{d-3} lacks the sequences coding for the first two domains but contains the rest of the coding sequence.

Comparison of HLA-B7, HLA-A2 and H-2K^b amino acid sequences shows that the third domain is the best conserved part

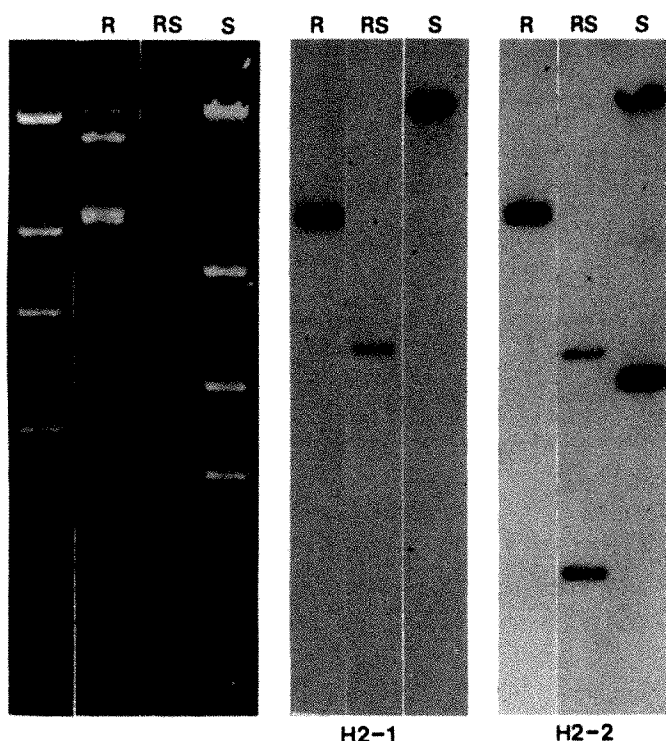


Fig. 1 Ethidium bromide-stained gel (left) and autoradiogram after transfer and hybridization with the H2-1 (middle) or the H2-2 (right) probes of λ HLA-12 DNA digested with *Eco*RI, *Sst*I and *Sst*I + *Eco*RI restriction enzymes. The first pattern on the left is *Hind*III-digested λ DNA giving fragments at 23.6, 9.46, 6.78, 4.34, 2.26 and 1.98 kilobases. Aliquots of digested λ HLA-12 DNA were run on a 0.5% agarose gel slab (overnight at 3 V cm⁻¹) and transferred to nitrocellulose according to Southern¹¹. The filter was baked at 80 °C for 2 h, pretreated for 4 h at 68 °C in 6 × SSC, 1 × Denhardt's mix¹² and hybridized overnight at 68 °C in 2 × SSC, 1 × Denhardt, 0.5 × SDS with H2-1 or H2-2 probes. The probes (*Pst*I-digested pH-2^d-3 DNA fragments) were recovered from a low gelling point agarose gel¹³, run on a small DEAE (DE52) column and labelled by nick-translation¹⁴ to specific activities of 1–4 × 10⁷ d.p.m. per μ g. After hybridization, filters were washed three times at 68 °C in 2 × SSC, 1 × Denhardt and 0.5 × SDS, twice at 20 °C in 2 × SSC, 50% formamide, 0.5% SDS, and twice in 2 × SSC. Exposure was 1–2 days at –70 °C with Ilford 'fast tungstate' screens. It is apparent from these patterns (see Fig. 2 for corresponding restriction map) that the H2-1 probe hybridizes only to the right of the last *Sst*I site in the insert whereas the H2-2 probe hybridizes to two sets of fragments which are in fact separated by ~1.9 kilobases.

of the molecule^{3,4}; the H-2^d amino acid sequence derived from the pH-2^d-3 DNA sequence shows 72.6% homology with that of HLA-B7 in the third domain⁷. These observations led us to speculate that there might be sufficient nucleotide sequence homology between H-2 and HLA DNA sequences to allow specific hybridization of H-2 probes to human HLA genomic segments.

Two DNA fragments were purified from the pH-2^d-3 plasmid after digestion with *Pst*I: one (abbreviated H2-1) is ~600 base pairs long and includes sequences coding for the transmembrane and cytoplasmic parts of the H-2 heavy chain and the other, H2-2, is 195 base pairs long and encodes most of the third domain⁷. Both fragments were labelled with ³²P by nick-translation and used as probes to screen a human gene library consisting of random DNA fragments recombined in λ Charon 4A, constructed by Lawn *et al.*⁸. About 10⁵ plaques were screened and several positive phages were purified. *In situ* hybridization was initially carried out with probe H2-1 at 63 °C

in 2 × SSC (rather than the usual 68 °C). However, all positive phages hybridized strongly with probe H2-2 at 68 °C. One recombinant, λ HLA-12, containing a 15-kilobase insert, was selected for further study.

λ HLA-12 DNA was analysed by digestion with *Eco*RI, *Sst*I and *Bgl*II restriction enzymes alone or in combination. The fragments were separated on agarose gels and analysed by Southern blotting and hybridization with the H2-1 and H2-2 probes. Figure 1 shows an example of such an analysis. The restriction map obtained for λ HLA-12 is shown in Fig. 2. H2-2 hybridized to two different *Sst*I fragments, which indicates that at least one cleavage site may lie within sequences homologous to this (third domain-specific) probe; H2-1 hybridized only to the right-ward *Sst*I fragment. Hybridization to the *Bgl*II and *Bgl*II + *Sst*I fragments confirmed these results, which indicate that the gene is oriented left to right on the map.

As *Sst*I was suspected to cleave within the region coding for the third domain, λ HLA-12 DNA was digested with *Sst*I restriction enzyme and labelled at the 5' or at the 3' ends. One of the fragments hybridizing with the H2-2 probe (see Fig. 1) was isolated after secondary digestion with *Eco*RI restriction enzyme and sequenced by the Gilbert–Maxam method⁹; labelling at the 3' and 5' ends allowed sequence determination on both strands from the same restriction site.

The DNA sequence obtained to the left of the second *Sst*I site in the insert is shown in Fig. 3. The protein sequence corresponding to a translation phase open in this region is shown above the nucleotide sequence; the HLA-B7 (ref. 3) sequence is shown on top. Figure 3 shows that the region sequenced corresponds to an HLA gene segment (amino acids 199 to 230 in

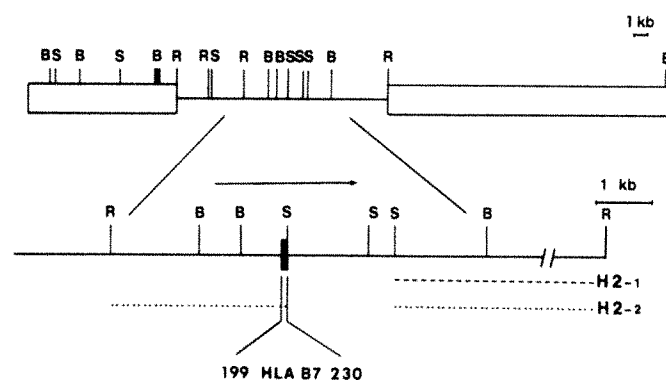


Fig. 2 Restriction map (R = *Eco*RI, S = *Sst*I, B = *Bgl*II restriction enzymes) of λ HLA-12. λ HLA-12 was identified and purified by three successive rounds of plaque hybridizations¹⁵ (1 × 10⁴, 3 × 10³ and 200 plaques per dish) using nick-translated¹⁴ (1–3 × 10⁸ c.p.m. per μ g) H2-1 or H2-2 probes (for conditions see Fig. 1 legend). Autoradiography was overnight at –70 °C using flashed films and Du Pont 'Cronex lightning plus' screens. λ HLA-12 DNA was prepared from a 600-ml culture of the *Escherichia coli* 1106 strain infected at a multiplicity of three. The phage was purified on caesium chloride step and continuous gradients and DNA was extracted using phenol, ether and ethanol precipitated. For mapping purposes aliquots (0.5 μ g) of the DNA were digested with restriction enzyme(s) and analysed in parallel on 0.5% agarose gel slabs. Southern hybridizations¹¹ were done as described in Fig. 1 legend. Top, complete map including the Charon 4A λ phage¹⁶ (with the right arm to the left to show the insert in the correct orientation with respect to gene sequences). Bottom, enlargement of the region in which DNA sequencing was performed, the direction of transcription is shown by the arrow. The dashed line indicates the *Eco*RI–*Sst*I fragment hybridizing with H2-1, the dotted line those hybridizing with H2-2. The area whose sequence is given in Fig. 3 is indicated by the box; the corresponding region in the HLA-B7 protein sequence is shown below with the numbers of the terminal amino acids. The approximate position for the third *Sst*I site in the insert is indicated.

HLA B7

199

A T L R C W A L G F Y P A E I T
A T L R C W A L G F Y P A E I T

5' GCCACCCCTGAGGTGCTGGGCCCTGGGCTTCTACCCTGCGGAGATCACACA

HLA B7

L T W Q R D G E D Q T Q D T E L
L T W Q R D G E D Q T Q D T E L

CTGACCTGGCAGCGGGATGGGGAGGACCAGACCCAGGACACGGAGACT(C) 3'

SsrI

We thank B. Cami for advice, help and materials, and F.M. Kourilsky for support and stimulating discussions. This work was supported in Marseille by CNRS and INSERM and in Paris by CNRS (ER 201, ATP 0061 and 1 15039), INSERM (SCN 20 and CRAT 72 79 104) and the Fondation pour la Recherche Médicale.

9. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
10. Orr, H. T., Lancet, D., Robb, R. J., Lopez De Castro, J. A. & Strominger, J. L. *Nature* **282**, 266 (1979).
11. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
12. Denhardt, D. *Biochem. biophys. Res. Commun.* **23**, 641-646 (1966).
13. Wieslander, L. *Analyt. Biochem.* **98**, 305-309 (1979).
14. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237-251 (1977).
15. Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
16. Blattner, F. *et al. Science* **196**, 161-169 (1977).
17. Tu, C.P.D. & Cohen, S. N. *Gene* **10**, 177-183 (1980).

Figures 1 and 2 show the results of such experiments performed with λ himA⁺, λ himAam79 (an amber mutation in the *himA* gene) or λ himA42 (a missense mutation). Labelled

1. Klein, J. in *The Major Histocompatibility System in Man and Animals* (ed. Gotzke, D.) 339-378 (Springer, Berlin, 1977).
2. Sveigaard, A. in *Immunology 1980, Progress in Immunology IV* (eds Dausset, J. & Fougereau, M.) 530-540 (Academic, London, 1980).
3. Strominger, J. L. in *Immunology 1980, Progress in Immunology IV* (eds Dausset, J. & Fougereau, M.) 541-554 (Academic, London, 1980).
4. Martinka, J. M. *et al. Biochemistry* (in the press).
5. Kvist, S. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
6. Ploegh, H. L., Orr, H. T. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6081-6085 (1980).
7. Bregegere, F. *et al. Nature* (submitted).
8. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157-1174 (1978).

extracts were subjected to two-dimensional gel electrophoresis. Figure 1a, b shows autoradiograms of gels of bacteria infected with λ himA⁺ and λ himAam79, respectively. The two autoradiograms are equivalent except for the absence, in the gel of the λ himAam79-infected cells, of an 11,000-molecular weight (MW) spot that is present in the gel of bacteria infected with λ himA⁺. Gels of extracts of uninfected cells were devoid of labelled spots (data not shown). As nonsense mutations are known to cause premature polypeptide chain termination, the absence of this spot tentatively identifies this 11,000-MW protein (marked *himA* in Fig. 1a) as the product of the *himA* gene (*phimA*). No protein that could correspond to the *phimAam79* amber fragment is seen. Note that the *phimA* spot seems to exhibit charge heterogeneity in the isoelectric focusing dimension. The separation of two apparent species of *phimA* varies somewhat with the conditions of electrophoresis.

Further evidence supporting the identification of the 11,000-MW protein as *phimA* comes from an additional experiment detailed in Fig. 2. In this experiment, the dose of UV rays was insufficient to eliminate completely all host protein labelling and a small amount of label was incorporated into ribosomal protein S10 which therefore serves as a convenient reference point (see

below for the location of S10 in this gel system). As in the previous experiment, cells infected with λ himAam79 (Fig. 2a) lack the 11,000-MW *phimA* protein that is made after infection with λ himA⁺ (Fig. 2b). Furthermore, the λ himA42 transducing phage directs the synthesis of a protein of identical molecular weight to *phimA* but with a more basic charge. This suggests that the *himA42* mutation has substituted an amino acid in the *phimA* protein that results in a net basic charge shift.

From the results of the experiments discussed above, we identify *phimA* as an 11,000-MW protein. The position of this protein in the isoelectric focusing indicates that *phimA* is a weakly basic protein²⁸. We do not know whether the charge heterogeneity of *phimA* reflects the state of this protein within the cell.

The λ himA transducing phage directs the synthesis of several proteins in addition to *phimA*, most of which have previously been identified by Springer *et al.*¹⁷ and Henneke *et al.*^{18,19}. These workers used the same transducing phage in similar experiments, but analysed the labelled proteins by one-dimensional polyacrylamide gel electrophoresis. We were thus able to identify the spots in two-dimensional autoradiograms by comparing their apparent molecular weights and relative intensities. In Fig.

Fig. 1 Autoradiograms of two-dimensional gels of UV ray-irradiated lysogenic bacteria infected with λ himA⁺ (a) or λ himAam79 (b). Logarithmic phase cultures of strain K447 (159*uvrA*⁻ carrying λ c⁺ind⁺Nam7, 53) growing at 39 °C in M56 media²⁶ plus 0.2% maltose at a concentration of 2×10^8 ml⁻¹ were pelleted and resuspended at a concentration of 2×10^8 ml⁻¹ in M56 media. Cells (10 ml) were irradiated with 1.6×10^4 ergs mm⁻² of UV light from a GE germicidal lamp (~2 min at a distance of 10 cm), pelleted, resuspended to the same concentration in M56 with 0.02 M MgSO₄ and incubated in the dark for 15 min with shaking at 37 °C. To 0.5 ml of these cells, phage were added at a MOI of 10, and adsorption was allowed on ice for 15 min. The infected cells were once again pelleted and the supernatant discarded. To initiate growth, 2.5 ml of pre-warmed M56, 0.2% maltose and 0.02 M MgSO₄ were added and the cultures shaken at 37 °C in the dark. After 5 min, 10 μ Ci of ¹⁴C mixed amino acids (55 mCi mmol⁻¹, NEN) were added and labelling was allowed to continue for 60 min. The cells were then washed with cold 0.01 M Tris pH 7.4 and 0.01 M MgCl₂ and prepared for isoelectric focusing (IF). Strain K37 and the λ himA transducing phage are described by Miller and Friedman¹². Phage stocks were prepared by lytic growth in K37. Lysates were concentrated by polyethylene glycol precipitation²⁷ and purified by isopycnic CsCl buoyant density centrifugation. The phage were dialysed against 0.01 M Tris pH 7.4 and 0.01 M MgSO₄. Two-dimensional non-equilibrium gels were run according to the method of O'Farrell²⁸. All reagents and reagents are described in ref. 16. The IF dimension was run for 1,400 V-h using LKB ampholines (pH 3.5–10). The second dimension was a 20% polyacrylamide SDS stacking gel system with a 23-mm stacking gel and a 90-mm separating gel. Second dimensions were run for 5 h at 3.3 W per gel. The gels were stained in 0.005% Coomassie brilliant blue in 0.01% trichloroacetic acid, 7.5% acetic acid and destained in 7.5% acetic acid. Most of the ampholines could be removed by first soaking the gel in stain made 50% with ethanol. The gels were hot vacuum dried on Biorad paper and arbitrary orientation marks were drawn with radioactive india ink to facilitate subsequent alignment of the gel with the autoradiograms. The gels were then exposed to Kodak X-Omat film. Spot identifications are described in the text. The arrow in b indicates the expected location of *phimA*. Exposure time was 60 d. Samples were prepared for IF in the following manner. The labelled cultures were pelleted in 1.5-ml Eppendorf tubes; 100 μ l of 0.01 M Tris pH 7.4 and 0.01 M MgCl₂ containing 100 μ g ml⁻¹ freshly prepared lysozyme (Worthington) were added and the pellet resuspended. The cells were freeze-thawed for five cycles in an ethanol-dry ice bath. Aliquots of 5 μ l of a 1 mg ml⁻¹ solution each of pancreatic DNase (Worthington) and RNase (Sigma) were added and incubated on ice for 10 min. Then, 50 mg of solid urea (Schwarz-Mann ultrapure) and 100 μ l of lysis buffer¹⁶ were added to complete the sample preparation. The extracts were stored at -80 °C. The molecular weights of proteins were estimated by comparing the mobilities of the proteins in the second (SDS polyacrylamide) dimension with the mobilities of known proteins: two internal standards whose MWs are known from primary structure determinations were IF3 (20,668)²⁹ and whose location was determined in this work; and *E. coli* HU protein (9,535)³⁰, identified with purified HU protein (from K. Abremski). An additional marker was ribosomal protein S10 taken as 10,500 (refs 31, 32).

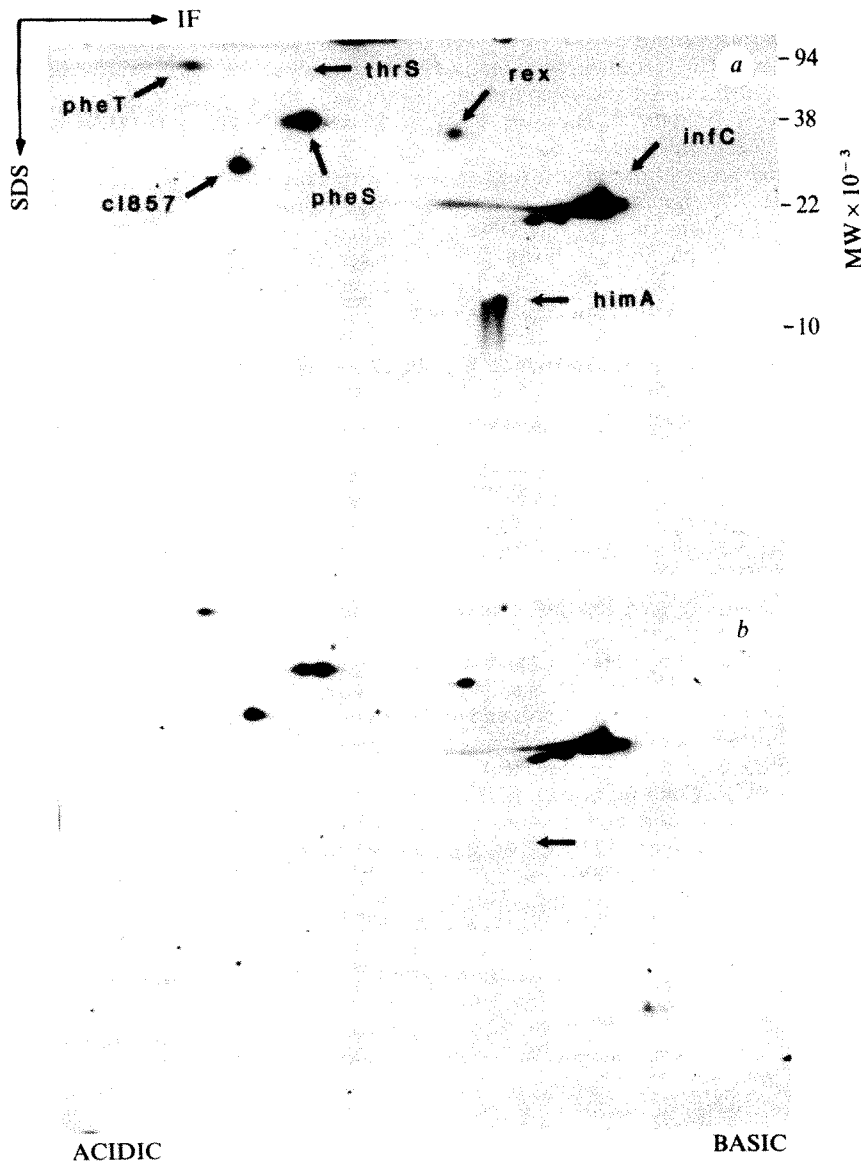
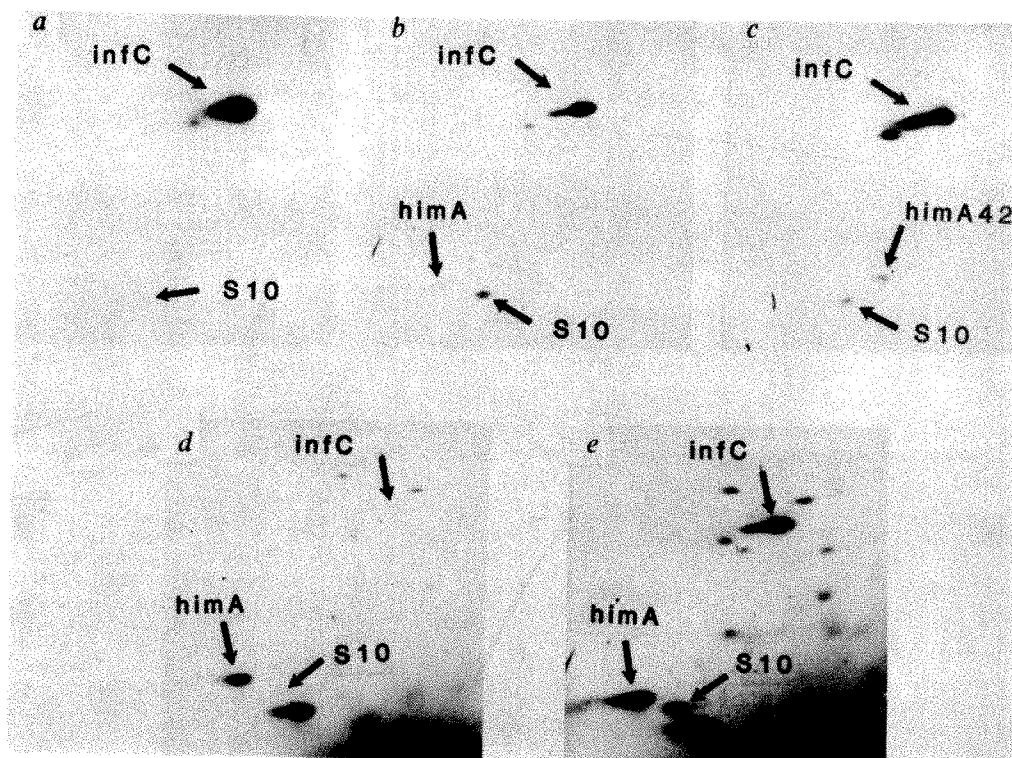


Fig. 2 Details of gels and autoradiograms of the regions surrounding S10 and IF3. *a-c*, UV-ray irradiated bacteria infected with λ himAam79 (*a*), λ himA⁺ (*b*) and λ himA42 (*c*). All three are autoradiograms. *d* Is the stained gel used to make the autoradiogram in *b* with the addition of purified IHF. *e* Is a composite overlay of *b* on *d*. Exposure time was 60 d. Purified integration host factor (IHF) was concentrated by precipitation with 20% trichloroacetic acid and resuspended in a 1:1 mixture of lysis buffer and 8 M urea.



1a, the spots are named for their respective genes. Thus, the spot marked *infC* is translation initiation factor IF3. As noted by Springer²⁰, the *infC* spot is the most heavily labelled and consists of a cluster of spots due to the different species of IF3²¹. We were able to locate the spot for IF3 in two-dimensional gels of total *E. coli* proteins by overlaying this autoradiogram on the stained gel. IF3 maps at coordinates (42, 54) on the standard non-equilibrium two-dimensional gel map of *E. coli* proteins of Bloch *et al.*²². The spots marked *thrS* (threonyl-tRNA synthetase), and *pheS* and *pheT* (the α - and β -subunits of phenylalanyl-tRNA synthetase) are also shown (the *pheT* spot does not resolve well in this system). Two additional spots have been identified as the products of the λ cI857 (repressor) and *rex* genes. These proteins, which are the only λ proteins synthesized in a repressed lysogen, were identified in separate experiments using non-transducing λ phage (H.I.M., in preparation).

Springer *et al.*¹⁷ have reported the presence of a fifth protein, p12, that they could not associate with a known gene. We have also found this protein in one-dimensional gels of our extracts. It is not *phimA* as it does not disappear when λ himAam79 is used for infection. (We were unable to detect *phimA* in the one-dimensional system, presumably due to the high background found in the region where *phimA* should migrate.) Furthermore, the p12 band does not resolve into a distinct protein spot in two-dimensional gels, indicating that it may not constitute a unique protein species as suggested by Hennecke *et al.*¹⁸.

The *E. coli* component required for integrative recombination has been extensively purified (H. A. N. and C. Robertson, in preparation). This protein (IHF) was assayed by its ability to complement purified Int protein²³ for *in vitro* recombination. The most highly purified material contains two polypeptides, and available evidence indicates that both are required for recombination activity. In the experiment shown in Fig. 3, 5 μ g of purified protein (fraction V) was mixed with a small amount of ¹⁴C-labelled total *E. coli* proteins (as reference proteins). Figure 3a shows the stained gel, *b* the autoradiogram and *c* a composite overlay of the autoradiogram and the stained gel. As shown in Fig. 3a, purified integration host factor (IHF) consists of two polypeptide species of molecular weights 11,000 (α , marked *himA*) and 9,500 (β , arrow). Both species exhibit apparent

charge heterogeneities in this gel system. The intensity of staining indicates an approximately equal molar amount of α and β . Note that the position of the 11,000-MW spot roughly corresponds to the position of *phimA* in Fig. 1a. In addition, both *phimA* and α exhibit similar patterns of heterogeneity.

Figure 3b shows the location of total *E. coli* protein labelled with ¹⁴C mixed amino acids. We have indicated the location of two reference proteins, IF3 (marked *infC*, see above) and ribosomal protein S10²². Note in Fig. 3b that although neither subunit is present in detectable quantities, S10 is positioned between the α - and β -subunits of IHF. Figure 2b, *d*, *e* shows the results of an experiment in which purified IHF was mixed with radiolabelled *phimA* (from a UV-ray irradiation experiment) and the mixture was subjected to gel electrophoresis. Only details of gels including the regions surrounding IF3 and S10 are shown. Figure 2d shows the stained gel, *b* the autoradiogram and *e* the composite overlay. The additional spots seen in Fig. 2d and *e* are unlabelled *E. coli* proteins derived from the extracts of the UV-ray irradiated bacteria. The radiolabelled *phimA* protein co-migrates with the purified α -subunits of IHF.

Thus, the *himA* genes codes for one of the subunits of IHF active in integrative recombination by λ DNA *in vitro*. This result fully explains the fact that strains mutant in the *himA* gene lack an active integration host factor and also eliminates the possibility that the *himA* gene influences λ integration solely at the level of synthesis of components of the recombinational apparatus. The demonstration that *phimA* is directly involved in integrative recombination has gained considerable importance in light of the recent finding that the *himA* gene also regulates λ integration at the level of synthesis of Int protein¹⁴. A feature of the transposon Tn3-encoded *tnpR* gene product is its involvement both in recombination and in control of transcription of transposon genes^{24,25}. Thus, dual function of site-specific recombination proteins may be a common characteristic.

The *E. coli* gene coding for the β -subunit of IHF has not been determined, although several candidates are available⁹ (H.I.M., in preparation). However, one or more of these gene products may be involved indirectly by regulating the synthesis or processing of either of the active subunits of IHF.

The fact that *phimA* is directly involved in recombination raises the possibility that *phimA* might also directly participate

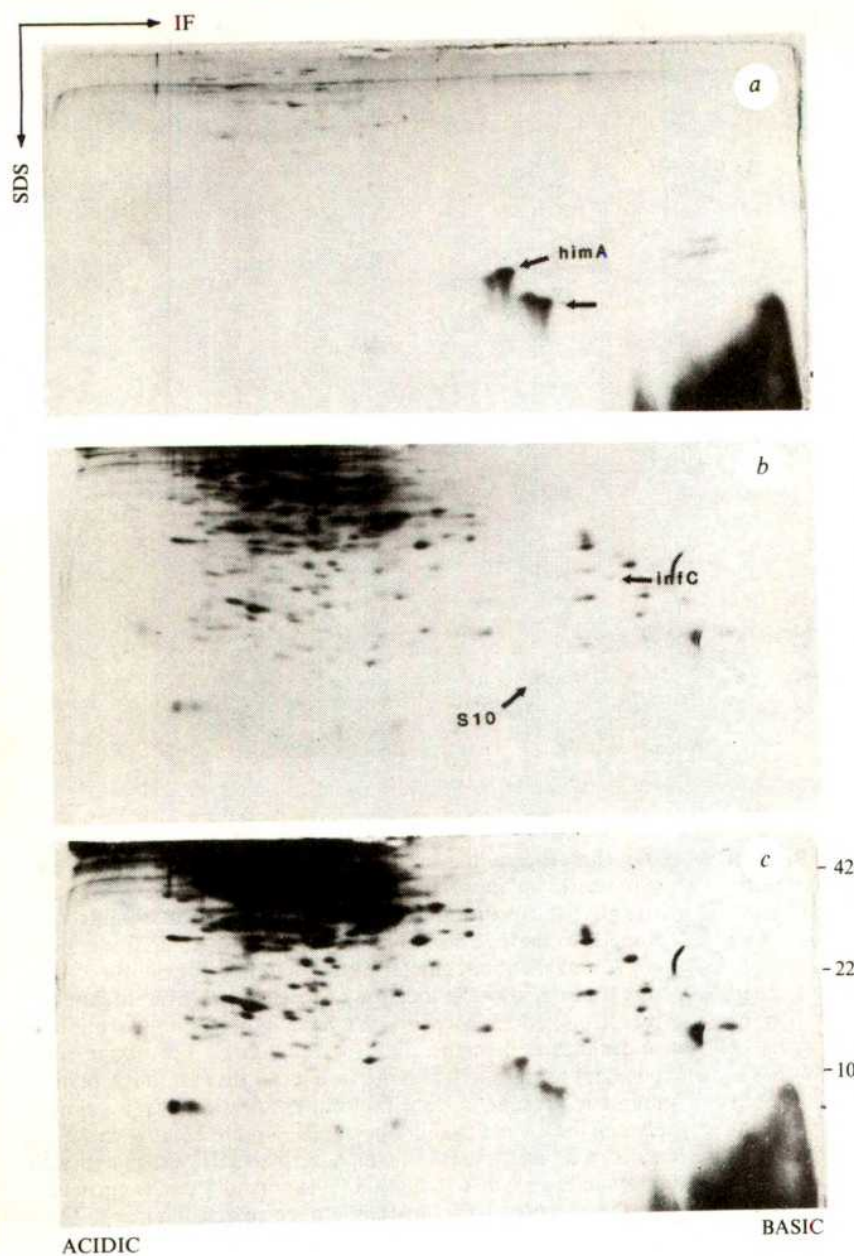


Fig. 3 Two-dimensional gel of purified integration host component. *a*, Coomassie blue stained gel. *b*, Autoradiogram of ^{14}C -labelled total *E. coli* proteins added as internal markers. *c*, Composite overlay of autoradiogram and stained gel. Additional spots in *a* are due to added *E. coli* proteins. Exposure was for 10 days. Dark staining material in bottom right of *a* is residual ampholines. For labelling bacterial cultures, $10\ \mu\text{Ci}$ of ^{14}C mixed amino acids were added to a logarithmic phase culture at a cell density of $2 \times 10^8\ \text{ml}^{-1}$ in M56 media and 0.2% glucose. After 2 min at 37°C , $100\ \mu\text{l}$ of 30% cassamino acid solution was added to the 2-ml culture and the cells were prepared for isoelectric focusing.

in several diverse processes known to be influenced by *himA* mutations, such as the growth of bacteriophage Mu and the precise excision of transposable antibiotic resistance elements. The availability of purified IHF should facilitate the biochemical investigation of these processes.

We thank Dr M. Belfort for the UV-ray irradiation procedure, M. Smith for introducing us to non-equilibrium gels, R. Steinberg for photographic displays, Drs S. Adhya, D. Court, D. Friedman and B. Paterson for encouragement and advice, H. Echols for critical reading of this manuscript, and Carolyn Walker for preparation of the manuscript. This work was supported by grants GM-17078 and 5T32-GM0727 from the National Institute of General Medical Sciences to H. Echols.

Received 20 October 1980; accepted 19 January 1981.

- Campbell, A. *Adv. Genet.* **11**, 101-145 (1962).
- Weisberg, R. A., Gottesman, S. & Gottesman, M. E. in *Comprehensive Virology* (eds Fraenkel-Conrat, H. & Wagner, R.) 197-258 (Plenum, New York, 1977).
- Nash, H. A. *Curr. Topics Microbiol. Immun.* **78**, 171-199 (1977).
- Kikuchi, Y. & Nash, H. A. *J. biol. Chem.* **253**, 7149-7157 (1978).
- Mizuuchi, K. & Nash, H. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3524-3528 (1976).
- Kotewicz, M., Chung, S., Takeda, Y. & Echols, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1511-1515 (1977).
- Zissler, J. *Virology* **31**, 189 (1967).
- Gingery, R. & Echols, H. *Proc. natn. Acad. Sci. U.S.A.* **58**, 1507-1514 (1967).
- Gottesman, M. E. & Yarmolinsky, M. B. *J. molec. Biol.* **31**, 487-505 (1968).
- Miller, H. I., Kikuchi, A., Nash, H. A., Weisberg, R. A. & Friedman, D. I. *Cold Spring Harb. Symp. quant. Biol.* **43**, 1121-1126 (1979).
- Williams, J. G. K., Wulff, D. L. & Nash, H. A. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 357-361 (Cold Spring Harbor Laboratory, New York, 1977).
- Miller, H. I. & Friedman, D. I. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 349-356 (Cold Spring Harbor Laboratory, New York, 1977).
- Miller, H. I. & Friedman, D. I. *Cell* **20**, 711-719 (1980).
- Miller, H. I. *et al. Cold Spring Harb. Symp. quant. Biol.* **45**, (in the press).
- Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **57**, 306-313 (1967).
- O'Farrell, P. H. *J. biol. Chem.* **250**, 4007-4021 (1975).
- Springer, M., Graffe, M. & Hennecke, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3970-3974 (1977).
- Hennecke, H., Springer, M. & Böck, A. *Molec. gen. Genet.* **152**, 205-210 (1977).
- Hennecke, H., Böck, A., Thomale, J. & Nass, G. *J. Bact.* **131**, 943-950 (1977).
- Springer, M., Graffe, M. & Grunberg-Manago, M. *Molec. gen. Genet.* **169**, 337-343 (1979).
- Suryanarayana, T. & Subramanian, A. *FEBS Lett.* **79**, 264-268 (1977).
- Bloch, P. L., Phillips, T. & Neidhardt, F. C. *J. Bact.* **141**, 1409-1420 (1980).
- Kikuchi, Y. & Nash, H. *Cold Spring Harb. Symp. quant. Biol.* **43**, 1099-1109 (1979).
- Heffron, F., McCarthy, B. J., Ohtsubo, H. & Ohtsubo, E. *Cell* **18**, 1153-1163 (1979).
- Arthur, A. & Sheratt, D. *Molec. gen. Genet.* **175**, 267-274 (1979).
- Monod, J., Cohen-Bazire, G. & Cohen, M. *Biochim. biophys. Acta* **7**, 585 (1951).
- Yamamoto, K. R., Alberts, B., Benzinger, R., Lawhorne, L. & Treiber, G. *Virology* **40**, 734-744 (1970).
- O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 1133-1142 (1977).

A non-disruptive technique for loading calcium buffers and indicators into cells

R. Y. Tsien

Physiological Laboratory, University of Cambridge, Cambridge CB2 3EG, UK

Present methods for measuring or buffering intracellular free calcium concentrations are almost entirely limited to robust and well anchored cells which can tolerate insertion of ion-selective microelectrodes or microinjection of calcium indicators or buffers into one cell at a time¹. A very few types of small cells can be loaded with buffers or indicators during controlled lysis²⁻⁴, but such procedures grossly perturb membrane integrity and soluble cytoplasmic constituents. Liposome fusion releases only trace quantities of the trapped solute into the cytoplasm^{5,6} and incorporates foreign lipid into the target cell membranes. I now describe a simple technique which loads Ca^{2+} -selective chelators into the cytoplasm of intact cells in suspension and avoids the disadvantages of previous methods. The chelators are made temporarily membrane permeable by masking their four carboxylates with special esterifying groups which then hydrolyse inside the cells, regenerating and trapping the original chelators. The method is demonstrated on red cells, mast cells and lymphocytes.

The Ca^{2+} -specific chelator used in initial trials (shown in Fig 1a) is a member of a homologous family^{7,8} of buffers and optical indicators that have unusual specificity for Ca^{2+} over Mg^{2+} , H^+ , Na^+ and K^+ . Its affinity constant for Ca^{2+} is $\sim 10^7 \text{ mol}^{-1}$, other members of the family offer both higher and lower affinities. The simplest esters (methyl or ethyl) of this chelator are not hydrolysed in human erythrocytes. Anhydrides, imidazolidines, or activated aryl esters were rejected because the end products would probably include chelator molecules irreversibly bound to intracellular nucleophiles. Trimethylsilyl esters were too rapidly hydrolysed extracellularly, while *t*-butyldimethylsilyl esters were potentially haemolytic. Success was finally achieved with the acetoxymethyl tetraester (Fig 1b), analogous acyloxy-alkyl esters are used to enhance intestinal absorption of antibiotics^{9,10}, and the accepted mechanism of hydrolysis is enzymatic cleavage of the acetyl ester linkage, followed by rapid spontaneous release of formaldehyde. Acetyl esterases are present in many cell types, as suggested by their ability to hydrolyse fluorescein diacetate^{11,12}.

The acetoxymethyl ester (Fig 1b) was prepared by dissolving the chelator free acid, 5 equivalents of dry ethyldisopropylamine (redistilled from reagent supplied by Aldrich Chemicals) and 6 equivalents of acetoxymethyl bromide¹³ in ethanol-free chloroform. After stirring for 24 h at room temperature, the solution was evaporated under vacuum and the residue dissolved in benzene, filtered to remove amine salts, and re-evaporated. Extraction of this residue with ether, followed by evaporation, yielded a soft gum which was largely but not entirely pure as determined by PMR spectroscopy. Solutions or

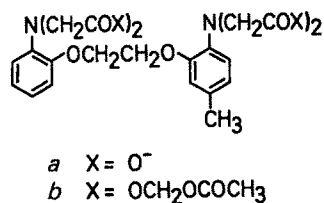


Fig. 1 Structures of the calcium chelator (a) and its acetoxymethyl tetraester (b). Hydrolysis of the ester should yield the chelator tetra-anion, four molecules each of acetate and formaldehyde, and eight protons.

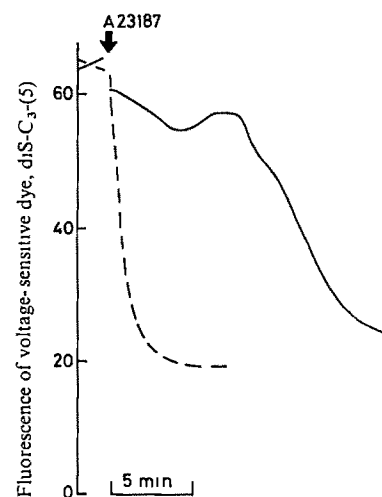


Fig. 2 Ability of intracellular chelator to delay the onset of the Ca^{2+} -induced K^+ permeability caused by A23187 (added as indicated to a final concentration of 200 nM). Abscissa, time, ordinate, fluorescence from 250 nM of the voltage-sensitive dye diS-C₃(5)^{18,19} in a suspension (0.04% haematocrit) of either normal human red cells (broken line) or cells loaded with chelator as described in the text (solid line). Fluorescence (excitation $630 \pm 10 \text{ nm}$, emission $665 \pm 10 \text{ nm}$) is expressed as a percentage of that seen in the absence of cells. The medium contained 130 mM NaCl, 10 mM KCl, 0.1 mM CaCl_2 and 10 mM tris(hydroxymethyl)-aminomethane-morpholinopropanesulphonate, pH 7.4, 22°C. Decreased dye fluorescence indicated membrane hyperpolarization due to increased K^+ permeability, resulting from increased $[\text{Ca}^{2+}]$. Ester treatment was followed by sufficient washing to remove essentially all acetate and formaldehyde. The low concentration of A23187 was adequate to induce K^+ permeability because of the absence of Mg^{2+} from the medium¹⁵. At an external $[\text{K}^+]$ of 10 mM, the effect of the dye on K^+ permeability may be disregarded¹⁷.

suspensions of the ester in physiological saline were prepared by injecting freshly made-up concentrated stock solutions in dimethyl sulphoxide into well stirred aqueous solutions.

Human erythrocytes were used in initial trials as they are readily available in large numbers, lack internal Ca^{2+} -sequestering organelles and proteins, and have a well characterized Ca^{2+} -triggered K^+ permeability. The low rate of endogenous H^+ production and the rapidity of transmembrane pH equilibration allow ester hydrolysis to be monitored by H^+ production in a pH-stat. There was no ester hydrolysis detected at neutral pH in the absence of cells, but with cells at 10% haematocrit, pH 7.4, 37°C, the partially purified ester (0.2 mg per ml of suspension) released protons at a rate decaying roughly exponentially with a time constant of $\sim 15 \text{ min}$. After 2 h, total acid production, corrected for endogenous H^+ production, was 1.56 mM, slightly more than expected from complete hydrolysis of perfectly pure ester. Cell lysis after 2 h was 5.5%, partly due to the continued vigorous stirring necessary in a pH-stat. Titration of an aliquot of the cell suspension with CaCl_2 and a Ca^{2+} -selective electrode¹⁴ showed that the amount of chelator accessible to the extracellular space was only $\sim 3\%$ of that expected from total ester hydrolysis. About this amount would be expected just from haemolysis during the loading procedure. However, once the cells were lysed by freezing and thawing, titration revealed $\sim 1.4 \text{ mmol}$ chelator per litre of cells ($\sim 90\%$ yield) with the correct Ca^{2+} affinity, within experimental uncertainty imposed by incomplete Ca^{2+} selectivity of the electrode at $p\text{Ca} \geq 7$. The amount of intracellular chelator did not significantly change either after resuspension of the loaded cells in chelator-free saline for 1 day, or after digestion with 0.1 M KOH for 2 days, showing that the chelator was well trapped and that hydrolysis was complete. Imperfect purity of the ester gum was probably a major reason why the yields of protons and chelator were not quite stoichiometric.

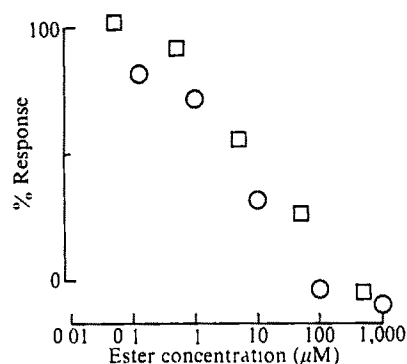


Fig. 3 Dose-response curve showing the ability of intracellular chelator to counteract the stimulation of histamine secretion from rat peritoneal mast cells by Con A plus lysophosphatidylserine. The cells were obtained and purified as described elsewhere²¹. The lysolipid was a gift of Dr G. A. Smith. The abscissa indicates the concentration of ester in which each batch of cells (10^5 per 25 μ l) was incubated for 30 min at 25 °C. The cell suspensions were then diluted to 500 μ l each, stimulated with Con A and lysolipid (final concentrations 100 μ g ml⁻¹ and 1 μ M respectively) for 15 min, and assayed for histamine release²¹. The ordinate indicates histamine release (corrected for spontaneous release) as a percentage of the stimulated histamine release in the absence of ester. This percentage is negative for ester concentrations ≥ 100 μ M because cells thus treated released less histamine after stimulation than the 'spontaneous release' controls that were neither stimulated nor treated with ester. Circles and squares represent two independent experiments in which spontaneous histamine release was, respectively, 2.9 and 5.0% of total cell histamine, and stimulation without ester gave releases of 23.6 and 11.0%. Control experiments showed that the inhibition by ester was not due to formaldehyde or acetate, as preincubation with up to 400 μ M formaldehyde (as would be produced by complete hydrolysis of 100 μ M ester) or 4 mM acetate showed either no effect on, or an actual enhancement of, stimulated histamine secretion.

Figure 2 shows an independent demonstration that these cells have trapped a cytoplasmic high-affinity Ca^{2+} chelator. In normal cells (broken line), addition of the Ca^{2+} ionophore A23187 causes an influx of Ca^{2+} , increasing the intracellular free calcium concentration ($[\text{Ca}^{2+}]_i$) and switching on a K^+ permeability^{2,15,16}. In medium containing 10 mM K^+ (ref. 17), this permeability hyperpolarizes the cells and can be detected as the decrease in fluorescence (ordinate) of a membrane potential-sensitive dye^{18,19}. Introduction of an intracellular Ca^{2+} chelator should mean that the Ca^{2+} influx has to continue much longer before the threshold $[\text{Ca}^{2+}]_i$ is reached. As expected, the ester-treated cells (solid line) show a marked delay before hyperpolarization develops, presumably reflecting the time required for the Ca^{2+} influx to titrate the trapped buffer.

Another cell type that accumulates chelator is the rat peritoneal mast cell, but the small number of cells and their internal complexity made use of the pH-stat and direct Ca^{2+} titrations impractical. Instead, chelator uptake was qualitatively assessed by the inhibition of $[\text{Ca}^{2+}]_i$ -dependent histamine release²⁰. Figure 3 shows that preincubation of mast cells with increasing concentrations of the ester progressively inhibited histamine release in response to concanavalin A (Con A) plus lysophosphatidylserine^{21,22}. Preincubation with ester was also able to stave off the normally immediate release of histamine induced by a moderate dose of A23187. However, higher concentrations of the ionophore or longer exposure times eventually overcame this inhibition and led to the release of most of the cellular histamine. Thus the effect of preincubation with ester is consistent neither with a blockage of Con A-induced Ca^{2+} permeability, nor with permanent fixation of the histamine granules or metabolic poisoning, but can be explained by a fixed quantity of high-affinity chelator preventing or delaying the rise of $[\text{Ca}^{2+}]_i$ to the threshold for secretion. No toxic effects have yet been

detected in red blood cells or mast cells from the formaldehyde which should be released by the hydrolysis of effective concentrations (up to several hundred μ M) of ester. This is probably because the high permeability of formaldehyde and the low cytotoxicity prevent its concentration exceeding millimolar amounts anywhere. Histological fixative concentrations are of the order of hundreds of millimolar, whereas for short-term exposures to several millimolar, mild and partially reversible pharmacological effects are reported^{23,24}. If formaldehyde toxicity presents difficulties, it should be possible to change the acyloxy-alkyl group to release alternative carbonyl compounds^{9,10}.

These experiments show how chelators trapped by ester hydrolysis can damp out short-term $[\text{Ca}^{2+}]_i$ transients and thus reveal qualitatively which cell functions are influenced by such transients. The method described here can be greatly extended by choosing a chelator which is also a sensitive optical indicator of $[\text{Ca}^{2+}]_i$. For example, a fluorescent methoxyquinoline (structure 3b of ref. 7) is readily trapped inside mouse thymocytes and pig mixed lymphocytes by its acetoxymethyl ester²⁵. At low indicator dose, the Ca^{2+} sensitivity of its fluorescence⁷ shows that resting $[\text{Ca}^{2+}]_i$ in lymphocytes is 123 ± 6 nM ($\pm$ s.e.m., $n = 27$), a much sought-after²⁶ but hitherto unmeasured value. Higher intracellular concentrations of chelator, up to at least 4.7 mM, are also easily attained by adjustment of cytochrome and initial ester concentration. Effects of such buffering dosages on capping and mitogenic stimulation are being studied, the fluorescence sensitivity allows quantification of the effectiveness of the buffer in damping changes in $[\text{Ca}^{2+}]_i$.

Intracellularly hydrolysable esters offer the first method for trapping membrane-impermeable chelators and indicators inside large populations of small intact cells without disrupting their membranes. A technically simple and reliable new approach to the measurement and control of $[\text{Ca}^{2+}]_i$, thus becomes possible. All cell types tested so far have proved able to accumulate significant amounts of chelator. As this technique has been shown to give the expected results in cell types where the role of intracellular Ca^{2+} is fairly well understood, it can now be applied with greater confidence to a wide variety of tissues where the function of Ca^{2+} is less well established. Analogous esters may also prove useful outside the field of Ca^{2+} function, for example to deliver and trap indicators of pH and free Mg^{2+} , chelators of toxic heavy metals, or other membrane-impermeable carboxylic acids.

I thank the SRC for financial support and T. J. Rink, V. L. Lew and G. A. Smith for helpful discussion.

Received 24 November 1980, accepted 27 January 1981

- 1 Ashley, C. C. & Campbell, A. K. (eds) *Detection and Measurement of Free Calcium Ions in Cells* (Elsevier, Amsterdam, 1979).
- 2 Simons, T. J. B. *J. Physiol., Lond.* **256**, 209–244 (1976).
- 3 Yingst, D. R. & Hoffman, J. F. *Biophys. J.* **23**, 463–471 (1978).
- 4 Baker, P. F. & Knight, D. E. *Nature* **276**, 620–622 (1978).
- 5 Hagins, W. A. & Yoshikami, S. in *Vertebrate Photoreception* (eds Barlow, H. B. & Fatt, P.) (Academic, London, 1977).
- 6 Campbell, A. K., Lea, T. J. & Ashley, C. C. in *Detection and Measurement of Free Calcium Ions in Cells* (eds Ashley, C. C. & Campbell, A. K.) (Elsevier, Amsterdam, 1979).
- 7 Tsen, R. Y. *Biochemistry* **19**, 2396–2404 (1980).
- 8 Tsen, R. Y. L. K. patent application 42927/78.
- 9 Jansen, A. B. A. & Russell, T. J. *J. chem. Soc.* **2127**–2132 (1965).
- 10 Ferres, H. *Chem. Ind.* **435**–440 (1980).
- 11 Rotman, B. & Papermaster, B. *Proc. natn. Acad. Sci. U.S.A.* **55**, 134–141 (1966).
- 12 Thomas, J. A., Buchsbaum, R. N., Zimniak, A. & Racker, E. *Biochemistry* **18**, 2210–2218 (1979).
- 13 Neuenschwander, M. & Iseli, R. *Helv. chim. Acta* **60**, 1061–1072 (1977).
- 14 Ammann, D., Gueggi, M., Pretsch, E. & Simon, W. *Anal. Lett.* **8**, 709–720 (1975).
- 15 Lew, V. L. & Ferreira, H. G. *Nature* **263**, 336–338 (1976).
- 16 Lew, V. L. & Ferreira, H. G. in *Current Topics in Membranes and Transport* (eds Bronner, F. & Kleinzeller, A.) (Academic, New York, 1978).
- 17 Simons, T. J. B. *J. Physiol., Lond.* **288**, 481–507 (1979).
- 18 Sims, P. J., Waggoner, A. S., Wang, C. H. & Hoffman, J. F. *Biochemistry* **13**, 3315–3330 (1974).
- 19 Hladky, S. B. & Rink, T. J. *J. Physiol., Lond.* **263**, 287–319 (1976).
- 20 Gomperts, B. D. in *Receptors and Recognition Series A*, Vol. 2 (eds Cuatrecasas, P. & Greaves, M. F.) (Chapman and Hall, London, 1976).
- 21 Smith, G. A., Hesketh, T. R., Plumb, R. W. & Metcalfe, J. C. *FEBS Lett.* **105**, 58–62 (1979).
- 22 Martin, T. W. & Lagunoff, D. *Nature* **279**, 250–252 (1979).
- 23 Fozzard, H. & Dominguez, G. *J. gen. Physiol.* **53**, 530–540 (1969).
- 24 Shrager, P. G., Strickholm, A. & Macey, R. I. *J. Cell Physiol.* **74**, 91–99 (1969).
- 25 Tsen, R. Y., Rink, T. J. & Pozzan, T. (in preparation).
- 26 Metcalfe, J. C., Pozzan, T., Smith, G. A. & Hesketh, T. R. *Biochem. Soc. Symp.* **45**, 1–26 (1980).

BOOK REVIEWS

Memories of the secret city

Rudolf Peierls

Los Alamos was the laboratory set up, during the Second World War, to design and produce atomic bombs from the fissile materials, uranium 235 and plutonium, provided by the gigantic installations of the "Manhattan District" of the US Army. A remote location in the hills of New Mexico was chosen partly because of the high degree of secrecy required, and partly because the work might involve hazardous experiments. Conceived originally as a settlement for 265 people — six scientists with their support staff and families — it had grown to many thousands by the end of the War, largely because the design of a plutonium bomb proved much more difficult than had been foreseen.

Much has been written about the achievement of the laboratory in solving the difficult technical problems so quickly, and about the life in this unusual community. The present collection of lectures, given at Santa Barbara in 1975, is a most interesting addition to this literature, containing the recollections of scientists, of scientists' wives and of an army officer.

Imagine a team of scientists, many of them at the top of their professions, working in improvised buildings, living in temporary housing, depending for their professional and domestic needs on the army, and subject to security restrictions they did not always see as necessary or consistent, imagine a collection of army personnel supporting the activities of scientists whose objectives they were not supposed to know, and whose methods and manners were quite foreign to the army's style of operation. There were bound to be conflicts, irritations and grievances. The recollections indeed include many examples of such friction. The fact that all major problems were somehow resolved, and the minor nuisances accepted, is a tribute to the good sense of the people involved, their feeling of urgency and in no small measure to the personality of the director, Robert Oppenheimer.

There is already some folklore about the place. The first lecture, by General Dudley, reports four versions of how Los Alamos came to be selected as the site of the laboratory — his own and three others — of which at least two are quite incompatible with his and with each other. The recollections of the residents do not show any such disagreements on facts, but they differ much in emphasis. Some seem to remember above all the irritations, the shortages, the beautiful school building which for a time had no qualified teachers, the strange facets of the voluntary mail

Reminiscences of Los Alamos, 1943–1945
Edited by Lawrence Badash, Joseph O Hirschfelder and Herbert P Broida Pp 188
(Reidel 1980) Hbk Dfl 50, pbk Dfl 20

censorship. Others stress the attractions of the landscape and climate, the contacts with the local Indian culture, and above all the proximity of a large number of unusually gifted and interesting people. But all would probably agree with Laura Fermi's remark:

One interesting thing to me is that during the war I was in Los Alamos only a year and a half, and *still* it seems such a big portion of my life. It was so different, it was such intense living, that it seems impossible that in my case it actually lasted hardly eighteen months.

Her account and those of Bernice Brode and Elsie MacMillan give a lively picture of the domestic scene. Feynman's "Los Alamos from Below" is as entertaining as one has come to expect from him. He reveals the secret of his notorious skill in opening combination locks, and how he accidentally stumbled on a safety hazard in a complicated plant design. There are hilarious accounts of his brushes with the censor and other bureaucratic elements. In stressing his "low" status he has his tongue in his cheek. He was recognized at once as a man of outstanding promise. The opening statement that he did not even have a degree, has to be withdrawn a little later when he admits that he did complete his PhD before going to Los Alamos.

Kistiakowsky's reminiscences form the most substantial part of the book. He

describes, in non-technical language, the work of his division in designing and making the "implosion" device. He and his collaborators, whose contributions are individually acknowledged, had to fight on two fronts: against professional engineers who insisted on using traditional methods for unprecedented problems, and against physicists who wanted to verify the functioning of the devices by clever, but ambitious and time-consuming, and not always reliable, tests. His lecture ends with a disarmingly frank account of how he had a little too much drink in celebrating VJ Day, and how he fired 21 high-explosive bursts in place of a 21-gun salute, only it came out to be 22.

John Manley and Edwin MacMillan describe the early stages in the creation of the laboratory. Joseph Hirschfelder contributes a brief outline of the work of the laboratory with thumbnail sketches of some of its most famous members. Norris Bradbury, director from 1945 to 1970, deals with the transition to a peacetime institution. There is a historical introduction by Lawrence Badash.

A few questions after each talk give the impression that the audience (or at least its more vocal members) followed the lectures less intelligently than the reader will, but they still elicited interesting answers.

All in all, this mixture of styles and approaches adds up to a lively picture of an unusual, and maybe unique, experience. □

Sir Rudolf Peierls is Emeritus Wykeham Professor of Physics at the University of Oxford. He spent 18 months working at Los Alamos during the Second World War.

One man's view of desertification

Peter D. Moore

Land, Man, and Sand. By James Walls Pp.336 (Macmillan, New York/Collier Macmillan, London 1980) \$19.95, £12.95

"WHAT? Not *another* world conference?" Barbara Ward exclaimed in *The Economist* on hearing of the proposed United Nations Conference on Desertification in 1977. "What? Not *another* book reporting the outcome of another world conference", we may be forgiven for retorting on being faced with *Land, Man, and Sand*, a summary of the Desertification Conference's findings.

It is, however, a slightly different approach to the need for disseminating the results of such conferences in that it is

written by a single author, and he an experienced writer. Instead of being presented with an uncoordinated, turgid set of papers, we have a unified, flowing account which draws liberally upon the original reports and documents. It is an interesting experiment in communication, the main disadvantage of which is that we are inevitably given one man's viewpoint, however fair and objective that man may attempt to be.

If there is one thing more boring than collections of conference reports, it is accounts of how conferences come to be convened. No doubt much thoughtful, bureaucratic effort goes into this area of scientific communication but, generally

speaking, it lacks the glamour or even scientific content which warrant an introductory chapter. The only vital information within it is that the core of the conference was to be comprised of 15 case studies of desert-spread research from widely separated parts of the world. It is these studies which have formed the basis for the remainder of this book.

There are two aspects of the character of James Walls which emerge from these chapters and which demand comment. The first is that he handles language skilfully, though some of his descriptive passages would be more appropriate in a novel than in a supposedly scientific text. The second is that he is eager to comment upon rather than simply to paraphrase the reports which form the basis for the book. His comments are generally critical, incisive, voluble and often completely lacking in any factual basis. He is a master of emotive writing and selective editing.

There is a danger here of falling into the same trap, so I will confine my specific comments to the one case study in which I have had first-hand experience, namely the Turan study of north-east Iran. Walls' report of this study contains a regular smattering of factual errors. He constantly equates the modern town of Shahroud with the Parthian capital Hecatompylos, more important, his reproduced maps confuse areas of moving sands with the saline plains, thus making a nonsense of subsequent discussion of the geography

and sociological effects of sand movements.

It is this kind of betrayal of his lack of knowledge of the places and the people in this study which leads one to take with a pinch of gypsum his hostile comments about the sociological base of the study. For example, the statement is made in the case study that the traditional pastoralist of Turan is concerned more with the condition of his animals than with that of the vegetation. This conclusion is regarded as vapid and patronizing by Walls. How he can level such criticisms without the necessary experience is difficult to discern. The sociologist/anthropologists involved in this study spent many months living with the people of Turan and spoke fluent Farsi. I would prefer to trust their judgement on the issue than the preconceptions voiced by the commentator.

Sadly, the book exhibits many such examples of the author expressing his opinion in areas where it is unlikely that he has experience to equal the originators of the case study reports. These reports incidentally, are available in complete form, free from editorial comment (held by the United Nations Environmental Programme) and I would strongly recommend interested parties to seek these for information rather than this book. □

Peter D Moore is Senior Lecturer in the Department of Plant Sciences, King's College, University of London.

Who's playing games?

Lindon J. Eaves

The IQ Game: A Methodological Inquiry into the Heredity-Environment Controversy By Howard F Taylor Pp 276 (Harvester/Rutgers University Press 1980) £18.95, \$19.50

The IQ Game arises from the author's "Sincere concern with the accurate assessment of the causes of variation in IQ." It is a detailed development of three principal criticisms of past research into the genetic basis of intelligence. The author's concern primarily with "heritability" and "IQ" may have limited his opportunity to explore in detail many other relevant aspects of research and theory in the general area of human quantitative genetics.

Dr Taylor's first two arguments have a familiar flavour to those who have followed the discussions of the past 15 years: (i) data are inadequate, (ii) parameters estimated from only one or two kinship correlations may be biased. The third main theme is an assault on the case where it is strongest — in the work of the path analytic and biometrical genetical schools whose long campaign to raise the standard of debate in the area is at last bearing fruit. Even here, where large numbers of kinship correlations are examined simultaneously with explicit models and statistical tests, the author suggests that the possibility of a very substantial environmental component cannot be excluded. In developing this position he relies heavily on a single model for the genetic and social transmission of intelligence suggested early in the 1970s by Morton and his associates, with its peculiarly idiosyncratic assumptions about the mating system and environmental transmission. The book depends also on the cogent criticisms and corrections of Professor Goldberger. It is convenient to have these reviewed in a more accessible form.

Dr Taylor shows that the remarkable consistency of multiple kinship data with a simple model for genetic and social transmission might be improved still further by the inclusion of numerous additional "environmental" components to explain anomalies in the data. Taken to the limit, indeed, it is possible to dispense entirely with a genetic component by recourse to as many *ad hoc* environmental parameters as there are kinship correlations to be explained. Those who play the "IQ game", according to Dr Taylor, make such arbitrary and implausible assumptions as will secure an estimate of heritability. This leads to the claim that a "plausible" relaxation of assumptions would reduce the estimated contribution of genetic factors and justify complete agnosticism about the causes of individual differences. In his analysis, however, relaxation of

James D. Watson THE DOUBLE HELIX

A NEW CRITICAL EDITION

Edited by Gunther S. Stent

The discovery in 1953 of the structure of DNA transformed the science of biochemistry. Written by one of the participants in this exciting scientific advance, *The Double Helix* has sold over a million copies since its publication thirteen years ago.

In this new critical edition the original text is accompanied by the reviews by leading scientists which greeted its publication, retrospective views of two other chief characters in the story — Francis Crick and Linus Pauling — and the original scientific papers in which the double helical structure of DNA was presented.

£10.00 ON SALE NOW

Weidenfeld & Nicolson

assumptions is not accompanied by any reduction in residual degrees of freedom. He is thus proceeding as if *any* set of environmental correlations is equally appealing.

It is here that the author's basic misconception of the "game" is revealed. However much the assumptions of the Hawaii model might be questioned on biological or psychological grounds, they are nevertheless founded upon a strong yet simple theory of genetic and social transmission. Such a theory comes close to predicting the observed distribution of individual differences in IQ. Science only degenerates into a "game" when additional components are imported *post hoc* with no clear basis in a theory of gene action or social interaction.

The provision of competing environmental theories which could match the classical genetic theory for simplicity and explanatory power is an exciting and challenging prospect. The book devotes three pages to a number of positive suggestions, of which the need to develop

environmental theories is but one. It is suggested that IQ should be treated as an index of an underlying construct rather than the actual phenotype of interest, that models for multiple cognitive measures be developed, that models and designs should be explored to allow separation of the developmental recursion of genetic and environmental factors.

Had these suggestions been published at the start of the past decade they would have been heralded as a positive and imaginative contribution. Now, however, Dr Taylor is not alone in his concern for truth. A substantial body of theoretical and empirical research in the wider area of human and behavioural genetics has been published in journals and conference proceedings. Over the past five years a significant start has been made on the study of precisely these issues and in response to those criticisms that the book develops in such detail. □

London J. Eaves is Fellow and Tutor in Experimental Psychology at Lady Margaret Hall and Lecturer in Experimental Psychology at the University of Oxford.

How microorganisms get their nitrogen

Ben Mifflin

Microorganisms and Nitrogen Sources: Transport and Utilization of Amino Acids, Peptides, Proteins, and Related Substrates. Edited by J. W. Payne. Pp 870. (Wiley 1980) £45, \$126.60.

THE lengthy sub-title of this book is intended, according to the editor, to provide a more complete indication of the coverage of the volume. However, even this does not really do justice to the number of topics covered. For example, at least four of the chapters deal with some aspect of nitrogen metabolism in higher plants, one of which is entirely devoted to the subject. Further, although the book set out to cover transmembrane metabolism, it also deals in some detail with the utilization of ammonia, amino acids and peptides. In all there are nearly 900 pages, spread over 26 contributions which are divided into four sections: Amino Acids, Peptides, Proteins, Related Systems. These have been written by a variety of international authors, many prominent in their field. Six chapters cover the transport and utilization of amino acids and peptides in bacteria, yeast and fungi. The genetics, regulation and energetics of the uptake of these nitrogen sources are dealt with separately in a further six chapters. Reflecting our greater knowledge of bacterial biochemistry, relative to that of other organisms, there is discussion of the membrane proteins responsible for amino acid transport. The section on proteins, although still dealing with transport, is more heavily biased towards utilization and re-utilization.

The final section of the book is to me what makes it really worthwhile, it is unusual in that the editor has included a number of contributions that should encourage the readers to broaden their horizons. Thus, there is a chapter on the γ -glutamyl cycle in animals and speculation as to how it might apply to microorganisms, and two on the utilization of nitrogenous compounds by green autotrophic plants. The utilization of ammonia and amides is well covered in three contributions and the reader is finally informed of some of the other functions of peptides, notably as antimicrobial agents and ionophores.

The contributions have generally been written with the aim of making them a readable and developmental account of the field rather than a combination of the author's latest results and his reference card library which, alas, all too often occurs in multi-authored volumes. The coverage of the literature is generally broad although not exhaustive (or exhausting) and, considering the size and number of authors, is reasonably up to date.

Whilst it is possible to argue with various points expressed and to pick out certain mistakes, this would be ungenerous to an editor and authors who have obviously striven hard to provide an informative and stimulating book, which I am sure will be a standard work in the field for many years to come. □

Ben Mifflin is Head of the Biochemistry Department at Rothamsted Experimental Station, Harpenden, Hertfordshire.

Complexity of melting

G. Gee

Crystal Melting. Macromolecular Physics, Vol 3. By Bernhard Wunderlich. Pp 363. (Academic 1980) \$42.50, £23.80.

THIS volume, the third in the author's treatise *Macromolecular Physics*, adds three chapters to the seven already published (in 1973 and 1976). The three volumes deal successively with the description of macromolecular crystals, with their growth and now with their melting. Although the author states that the first part of the project is now complete the next volume is evidently planned in some detail, for there are numerous references forward to it for thermodynamic discussions which are needed in this volume. This spreading of intimately interconnected topics over four volumes and ten years is a necessary consequence of a single author undertaking such a huge task, but it is not very helpful to today's reader to be told (p. 104) "the derivation is obvious by comparison with the results listed in 11.2.3" (Vol. 4).

In any elementary thermodynamics course melting is presented as a simple example of equilibrium between two bulk phases: a perfect crystal and its melt. Why then a whole volume on melting? In the first place, polymer crystals are characteristically small and imperfect. 100 per cent crystallinity is seldom attainable. Even the "melt" may require further specification, for example molecular weight and perhaps cross-linking or (in copolymers) composition. Finally, equilibrium between the two is in most cases almost impossible to achieve. These problems are recognized in the chapter divisions: equilibrium melting, irreversible melting, copolymer and isomer melting. In each chapter a discussion of general principles is followed by a detailed and critical survey of what is considered to be the most important and reliable experimental work.

This volume, like its predecessors, is addressed primarily to the specialist reader. He can safely be assumed to have at least a general knowledge of the topics discussed in other volumes (including those yet to appear). For him it is a splendid book, a reliable guide to the literature, offering a considered judgement but quick to point out difficulties. It is altogether too detailed for a reader wanting only a general account of this aspect of polymer physics. Between the two extremes will be many, engaged perhaps in the making or application of polymers, who will find it a useful source of reference and comment on the behaviour of a particular material. □

G. Gee is Emeritus Professor of Chemistry at the University of Manchester.

The second edition of Gregory Bateson's *Naven*, the classic study of the culture of the Iatmul people of New Guinea, has been reissued by Wildwood House, price £4.50.

Aspects of the oceanography of shelf seas

Henry Charnock

The North-West European Shelf Seas The Sea Bed and the Sea in Motion. Vol II, Physical and Chemical Oceanography, and Physical Resources Edited by F T Banner, M B Collins and K S Massie Pp 337 (Elsevier Scientific 1980) Dfl 185, \$90.25

IN THIS addition to the Elsevier Oceanography Series, five specialists deal with the tides, the wave climatology and with the physical oceanography of the Irish Sea, the North Sea, and the Celtic Sea and English Channel. One chapter deals with tidal power and a further seven are devoted to trace metals, to artificial radioisotopes, to sediment transport and to chemical and seabed resources (other than hydrocarbons). Volume I of this two-volume set dealt with the geology and sedimentology of the same area.

The physical oceanography contributions (Irish Sea by Bowden, Celtic Sea and English Channel by Pingree, and North Sea by Lee) are complementary, treating the same types of phenomena from different points of view. Coastal oceanography is clearly more complicated, if not more difficult, than that of the deep ocean conditions are much less homogeneous so that it is difficult to study idealized processes. But the economic and aesthetic

importance of some of the problems has led to an increase in the effort being put into trying to solve them. The numerical model succinctly described by Pingree should prove an increasingly valuable tool. Similar models are referred to by Huntley in his scholarly chapter on tides.

Of the other chapters, those on wave climatology (Draper), trace metals (Burton, Young) and artificial radioisotopes (Mauchline) are essentially factual. It is especially in these chapters that the delay between the preparation of the book and its publication is to be regretted, new measurements are being made, and old results updated, at a gratifyingly rapid rate.

Contributions on sediment transport and seabed resources (both by Banner) and the extraction of inorganic materials (Massie) illustrate the possibilities for exploitation of the ocean. They should be read in the context of the other chapters and in particular with that on socio-

economic development and the management of coastal zones, by Harding, which gives a good account of how difficult it is to use the ocean wisely, even (especially?) in a confined region bounded by developed nations. The north-west European shelf seas (and perhaps especially the North Sea) represent a microcosm of what is increasingly recognized as a global problem of the future.

Since this region is one of the relatively few with large tidal range, it was perhaps sensible to include a chapter on tidal power (by Wilson). This gives a clear account of the Rance Scheme, of the USSR experience and of the UK prospects, but could be better appreciated if there was a related account of wave energy, and perhaps of wind power.

There is, one supposes, rather little in this book which could not be found elsewhere, but it is convenient to have the material gathered together into one volume. □

Henry Charnock is Professor of Physical Oceanography at the University of Southampton

Mycotoxin biosynthesis reviewed

E. Arthur Bell

The Biosynthesis of Mycotoxins A Study in Secondary Metabolism Edited by Pieter S Steyn Pp 432 (Academic 1980) \$44, \$24.80

WIDESPREAD interest in the mycotoxins dates from the discovery of the aflatoxins in the early 1960s. Since then much has been published on these compounds and, as Sir Derek Barton states in his foreword to this volume, "the time is certainly ripe for an authoritative book". The present "authoritative book" is a collection of chapters written by leading workers in the field.

In the first chapter Bu'Lock discusses mycotoxins as secondary metabolites derived variously from polyketides, peptides, terpenes and other precursors. He emphasizes, moreover, that secondary metabolism, including mycotoxin biosynthesis, can be affected dramatically not only by the genetic modification of a fungus but also by the conditions under which the organism is grown or has been grown in the past. In succeeding contributions specific groups of mycotoxins are dealt with in detail. The biosynthesis of lysergic acid and the ergolines in *Claviceps* species is the subject of a contribution by Floss and Anderson. The formation of tricothecenes from mevalonate via farnesyl pyrophosphate in the Fungi Imperfecti is discussed by Tamm and Breitenstein, while Steyn, Vleggaar and Wessels describe the various techniques which have been used to elucidate the pathways of aflatoxin

synthesis from polyketide precursors.

The ergochromes, also of polyketide origin, are dealt with by Franck, while Yamazaki provides us with information concerning a heterogeneous collection of neurotropic mycotoxins, including aspergillid acid, which is formed from leucine and isoleucine, and slaphramine, a lysine derivative.

The biosynthesis of patulin and penicillic acid is considered by Zamir and that of the cytochalasins, a group of mixed biosynthetic origin, by Tamm. Kirby and Robins write of gliotoxin and related compounds, Holzapfel of cyclopiazonic acid, and Sankawa of the anthraquinonoid mycotoxins. The final chapter by Vleggaar and Steyn on "miscellaneous mycotoxins" provides even more evidence of the diversity of the biochemical pathways which lead to these fungal products.

It is a pity that the reader is given so little information about the physiological effects produced by these compounds and the nature of the organisms to which they are toxic. Biosynthesis is, however, the proclaimed subject of the book and the editor and authors have provided us with a veritable mine of information on the subject. This book will prove invaluable to all chemists, biochemists and biologists interested either in secondary metabolism or in the toxins of the fungi. □

E. Arthur Bell is Professor of Biology at King's College, University of London

**Can the U.S. stop
advancing when
Russia is advancing
nuclear power?**

**FROM SCIENTIFIC SEARCH
TO ATOMIC INDUSTRY**

By A. M. Petrosyants

388 pages 6 x 9 Hardbound \$17.90

This is a historical reference book to make the USSR's nuclear science and nuclear power development understood. Some pro-Russia interpretation is evident, but it gives the objective reader on the USSR, understanding and a comparison to USA technology.

Please send me _____ copies of
FROM SCIENTIFIC SEARCH
TO ATOMIC INDUSTRY
@ \$17.90 each TOTAL COST \$
Payment in full enclosed

Make check payable to
AMERICAN NUCLEAR SOCIETY
555 N. Kensington Ave. DEPT 5
La Grange Park, Illinois 60525

Circle No 14 on Reader Enquiry Card

16 April 1981

Reagan must live with arms control

Arms control has fallen off the agenda in Washington since the accession of Mr Ronald Reagan on 20 January, and nobody is surprised. Mr Reagan himself has a consistent record of indifference on the subject, now strengthened by the Administration's anxiety to be seen to be tough with the Soviet Union. The new Secretaries of State and of Defense, Mr Alexander Haig and Mr Caspar Weinberger, are unlikely to be urging that the Administration should push for further measures of arms control. In this climate, it is therefore a sign of grace — and of European and especially German influence in Washington — that the bilateral talks with the Soviet Union on the problems of limiting "theatre nuclear weapons" are continuing. Yet arms control cannot remain indefinitely in limbo. Sooner rather than later, perhaps in the last quarter of the year, the Administration will be forced to talk to the Soviet Union about something, European interests that there should be substantial progress in this direction are too strong indefinitely to be ignored. Then the non-nuclear signatories of the Non-Proliferation Treaty who went away grumbling from the review conference last September (and whose silence since is mystifying) will become thorns in the flesh of the superpowers again if nothing is done "to undertake effective measures in the direction of nuclear disarmament" as promised by the treaty. Whatever its inclinations, the Administration will find it has no choice but to talk. Where should it begin?

It should start with domestic housekeeping. As things are, the United States is better prepared for building nuclear weapons than for agreeing to dismantle them. To be sure, there is a dedicated government agency called the Arms Control and Disarmament Agency, which in its time has done sterling service and which has also been a means by which distinguished outsiders could make a mark on United States policy. The trouble is that the agency is only one of several sources of analysis and opinion on arms control proposals licensed to speak their minds in Washington. The State Department has often behaved as if the agency sheltering under its wing was a refuge for hopelessly idealistic liberals. But even the State Department cannot expect to have the last word. The Department of Defense and the Chiefs of Staff are separately entitled to speak up, while the White House itself in normal times is well equipped to disagree with everything that everybody else has said. The result is that proposals for arms control treaties usually find their way to the US Senate encumbered by a bewildering variety of conflicting opinions from within the Administration. Since 20 January, however, the Administration has been even less well prepared for action on arms control. The arms control agency has been without a director, and it is not yet clear whether the proposed nomination of Mr Eugene Rostow, a super-hawk, is intended to enable the agency to stand up to Mr Haig or, alternatively, to emasculate it (Mr Rostow once told the Senate Foreign Relations Committee that the only agreement between the United States and the "Russians" of which he thoroughly approved was that for the purchase of Alaska).

Even at this late stage, the Administration would be well advised to rebuild this cumbersome apparatus. The arms control agency owes its existence to the conviction, in the early 1960s, that if arms control were left to the Pentagon and the State Department, nothing would ever be accomplished. It is, however, folly to suppose that the agency could carry through the Senate agreements on arms control to which the Pentagon and the State Department were opposed. Then, with the passage of time, it has

become more and more widely recognized, even by the hawks, that arms control is an integral part of foreign policy and of national security. So much is clear from the working of the healthy clutch of international treaties accumulated since the ban on testing nuclear weapons in the atmosphere signed in 1963. Salt I, which limits the deployment of anti-ballistic missile systems, is more than a mere token of self-restraint by the superpowers. Because of the meetings of the standing consultative commission set up to supervise the working of the treaty, it has become a valuable means of exchanging data and opinions on technical issues. This, for example, is how the function of the Soviet "Backfire" intermediate bomber has been understood in Western Europe and the United States. So, as things have turned out, the process of arms control would be strengthened if the arms control agency were now firmly integrated with the State Department, not left on a limb where it can be disowned whenever the State Department and the Pentagon think it has gone too far. The overriding need is that the Administration should have a mechanism for arriving at one view, not half a dozen, of what it seeks from negotiations.

The Administration must also decide what it is prepared to talk about. The conventional wisdom is that Salt II, the agreement limiting the numbers of Soviet and American strategic weapons that was signed nearly two years ago but never ratified, is now dead. The Soviet occupation of Afghanistan was the occasion (in an election year) for President Carter tactfully not to press the treaty on an unwilling Senate. But nobody is sure of the status of a treaty which is signed but not ratified. In other fields — the Threshold Test-Ban Treaty, for example, which was signed in 1974 but never ratified — both superpowers behave as if the treaty had the force of international law, and complain if they think that one or other of them has violated the agreement. The chances are that both the United States and Soviet governments will find it prudent to behave as if Salt II were a reality. And long before 1985, when the unratified treaty would normally expire, those now saying that Salt II is dead will be crying "Long live Salt III". They will be referring to an amalgam of Salt II and whatever agreement emerges from the talks now under way on the technically daunting problem of limiting nuclear delivery systems of less than strategic range — cruise missiles, Pershing II and, for that matter, the "Backfire" bomber. Salt II is in many ways an unsatisfactory treaty (that could be improved by further negotiation) but the risks of letting it lapse will, long before 1985, seem too great for any administration to contemplate lightly.

The other dead duck from last year's agenda is the negotiation of the Comprehensive Test-Ban Treaty. A year ago, the United States Administration was saying that two months hard talking, and a sufficient measure of political will, would see the treaty signed and sealed. Since the arrival of Mr Reagan and his colleagues, that cheerful talk has quite evaporated. Instead, the military and the weapons laboratories have become eloquent about the risks of dispensing with weapons tests. The problem they identify is real enough, and comes in three parts. First, there are problems of reliability. Nuclear warheads deteriorate in performance with the passage of time, some of them because components such as tritium have relatively short half-lives, some of them because more conventional chemical components go off. The result is that some types of warheads must be refabricated at regular intervals and, in an ideal military world, tested. But there is no reason why the rebuilt bombs should not follow the original designs exactly. Second, there are problems of development. If

somebody designs a new delivery system, should not be weapons laboratories be asked to tailor the ideal warhead to it? The weapons laboratories answer enthusiastically and affirmatively. The truth, however, is that both the superpowers must now have such a range of efficient warheads that the search for perfection cannot substantially affect national security. Third, more important but less tangible, the military (on each side) would be psychologically uneasy (and the weapons laboratories would lose some of their best people) if testing were prevented by means of a comprehensive test-ban treaty. The strength of this objection was consistently underestimated in President Carter's unrealistic times. The result is that too little has been done, in the United States and elsewhere, to prepare the military for living without tests. It would be too much to hope that Mr Weinberger will take on this task so early in his career at the Pentagon. The best hope now is that the pressure of events will persuade him or his colleagues of the importance of the exercise.

For the immediate future, then, the prospect for substantial progress in arms control is far from bright. The new Administration will find itself dragged reluctantly towards Salt III. The Comprehensive Test-Ban Treaty will hang fire until opinion changes towards the second half of Mr Reagan's term. Wrongly, the Reagan Administration will be pilloried when it sets

about the dismantling of President Carter's clumsy anti-proliferation apparatus and may even be persuaded as a result that it had better stay clear of anything to do with arms control. That would be a great misfortune, especially because the agenda the Administration has inherited includes two treaties whose ratification by the United States Senate would bring benefits that are totally unalloyed. The Threshold Test-Ban Treaty (limiting tests to the equivalent of 150,000 tonnes of TNT of explosive power) includes provision for the definition of allowed testing sites, and for the pre-arranged calibration of the seismic characteristics of the underlying rocks, that would immediately improve the efficiency of the seismic detection networks on both sides. The treaty to regulate the use of peaceful nuclear explosions (with which even the Soviet Union now seems disenchanted) provides for similarly open exchanges of technical data about the characteristics of nuclear explosions. Would it not make sense, even for an administration anxious to demonstrate its firmness with the Soviet Union, to ratify (with the assistance of the Senate) these treaties? For as things are, the United States government (like that of the Soviet Union) dares not break the rules agreed, but cannot enjoy the benefits that would accrue from ratifying the treaties. Even Mr Reagan and his colleagues must know that that is absurd.

Can British education be reformed?

The British educational system, long and widely known for its waywardness (the encouragement of school students to abandon those studies of which they are most in need) and excesses (the preoccupation with examinations), may be about to change. The government has now published a document, *The School Curriculum*, that sets out a recipe for a uniform and rational curriculum for British schools and which also promises (some teachers say "threatens") a continuing interest in progress towards its modest goals. Others than the British may not fully appreciate the momentous and historic significance of what the British government has done. Although the education of the young is financed largely from taxation, the central government has no constitutional right to intervene in what is taught in British schools. The only legislative constraint on the school curriculum, embodied in the Education Act of 1944, is that publicly supported schools must provide religious instruction of some kind. For the rest, local educational authorities are constitutionally the final arbiters of the curriculum, but it has never been finally decided whether even they can tell head teachers how the school curriculum should be designed. Indeed, head teachers themselves are frequently at loggerheads with teachers and teachers' unions because their power to decide what should be taught in their schools is circumscribed by teachers' insistence on "academic freedom". Throughout the 1960s, when schools were sufficiently well supplied with funds to be able to experiment with the curriculum, successive central governments justified their cowardly detachment from what has happened by saying that they had no authority to intervene. It is therefore all the more creditable that central government should have chosen to intervene at this point, when there are no substantial funds with which to grease the palms of recalcitrant education authorities.

It remains to be seen how effective the new interventionist regime will be. *The School Curriculum* has only one important argument — that there should be a core curriculum in all publicly supported schools, consisting of language (usually English language), mathematics and science, and that no student in a secondary school should be allowed (let alone encouraged) to give up such a study before the age of 16. In addition, the document says, students in secondary schools "should undertake some study of the humanities designed to yield lasting benefit". The document recites predictable platitudes about the importance of moral education, preparation for adulthood and the like, and is intelligent but also defeatist about foreign language learning. Must not such a moderate and modest programme of reform

surely succeed?

Unfortunately, the outcome is by no means assured. Schools and the teachers' unions have been quick to say that the government cannot will the ends without providing the means — and that there is not merely a shortage of money but of people. For decades, the British educational system has been crippled by the lack of confident teachers of the essential parts of the curriculum as now defined — science, mathematics and modern languages. Even in first language teaching, trendily dubbed "communication skills" by too many schools, the efforts of able teachers are too often offset by those of teachers who believe that any string of words is a measure of a student's creativity, and therefore to be valued. The British government must know that even its modest proposals will not have the influence expected of them unless it can arrange for a general improvement in the quality of teaching. But recruiting more specialists, people with good degrees in mathematics, physics or even French, is no solution. A degree in mathematics is no assurance that a person will be able to explain the working of a slide-rule to a class of 12-year-olds. Nor will similar qualifications in other fields help primary school teachers to persuade even younger children that they are genuinely teachers.

Having embarked on the courageous course of pretending to have an opinion on the school curriculum, the British government has in the process saddled itself with an administrative problem. How will it enable the schools, at least some of which are eager to follow the instructions now implicitly handed down, to do what is asked of them? The most immediate need is not (as the teachers' unions and many schools have been saying) to increase the supply of specialist teachers but to rid the schools system of the tyranny of the public examinations that now dominate the curriculum. Fortunately, having agreed that there should be a reorganization of the public examinations provided for 16-year-olds, the government is well placed to do just this. The ideal would be that it should let schools (or their education authorities) provide school-leavers with suitable certificates and that the public examinations for 18-year-olds, the most distorting of present influences on the school curriculum because of their use as means of selecting students for higher education, should be abolished. Then, paradoxically, the government would find that, having made its point about the pattern of education, it could confidently allow schools to get on with the only task for which they are equipped — teaching young people. Left to themselves, but with clear objectives, they would find the teachers that they need.

European biotechnology lies in disarray

Commission's plan lacks friends

Trouble looms again for the ill-fated European research programme in biomolecular engineering, conceived in 1976 and, after much consultation, proposed to Community states in final form by the European Commission last summer. The Council of Ministers, which has the final authority over such proposals, placed the £16 million (26 million European Units of account (EUA)) five-year programme with its research committee for comment. And the committee now says that it cannot agree on the content, scale or cost of the programme.

The proposal was innovative in its time — it was praised by Jacques Monod and has been broadly backed by European industry, but it now looks unlikely to get approval in any form before November, the next date for a Council of Community Research Ministers. If the research committee had agreed, there was a chance that the programme would have gone through as a "point A" — a proposal that can be approved by any council without debate.

The programme deals with potential scientific bottlenecks in enzyme and recombinant DNA technology — such as provision of cofactors in enzyme-induced synthesis, or problems of expression efficiency in genetically engineered bacteria — through "indirect actions", where the Commission selects and organizes groups around Europe to investigate specific questions and pays half the costs (see *Nature* 283, 125, 1980).

Originally the programme contained six projects, but discussions last year within CREST, a committee of government representatives in which the Commission takes soundings of national views, reduced the number to four, eliminating areas with likely immediate medical or industrial application, and trimmed the budget to 15 million EUA. French representatives requested that half this sum be spent on education and training.

Eventually a "CREST compromise" was worked out, with 4 projects and 15 million EUA of which 20 per cent would be for education. But, at the level of the research committee of council, this compromise has proved unacceptable — as has a further compromise proposed by the secretary of the committee which would reduce the spending still further to 11.8 million EUA. The French members want their 50 per cent education and German members want the number of research projects cut to two, restricting the programme to agricultural applications (basically gene transfer in plants) and

safety questions. Britain appears to back the CREST compromise, or something close to it, all other states would have preferred the original proposal but are prepared to accept that of CREST.

In fact the proposal has fallen foul of the two major obstacles to Brussels-initiated research. First, the concept of a "European" industry benefiting the Community as a whole has not yet taken root, and if a piece of research has a potentially large economic benefit — such as research into cloning vectors, one of the original six proposals — nations are unwilling to share their expertise. The French view of the biomolecular engineering programme exemplifies this attitude: the government wishes the programme to be reduced to a training programme in

recombinant DNA techniques (it has had trouble recruiting for its genetic engineering company, Transgene) and effectively rejects the idea of joint research.

The second objection to Brussels-based work is that if nations already have a research programme in a particular area, they are unlikely to set aside funds in that area for work undertaken outside their control. This is broadly the position of German biotechnologists: they have a large national programme, but research money is tight, and contributing to a Commission-based programme would only distort their priorities.

More Community-minded biotechnologists in both France and Germany, however, point out that long-term progress

German physics in row about Einstein

The German Physical Society (Deutsche Physikalische Gesellschaft) has been shaken by a fierce row over an attack on Einstein published last November in *Physikalische Blätter*, the society's membership journal. Now the editorial board of the society, and the president of the society, Professor Horst Rollnik of the University of Bonn, have dissociated themselves from the original article, by Professor Albrecht Unsöld of the University of Kiel. In a statement in the March issue of *Physikalische Blätter*, they say that Unsöld's article contains statements that are open to misinterpretation, or even false, and promise that the journal will carry an article by an historian of science dealing with Einstein's links with Germany. There is, however, no suggestion that the journal will publish the critical letters it has received in the past few months.

Unsöld, a theoretician now 75, appears to have begun his attack on Einstein at a symposium held at the University of last May, when he said that Einstein (and Haber) was guilty of crimes "no less serious than those of Hitler".

Unsöld says that Einstein's earliest papers were deficient in references to the work of other physicists because of his "narcissism". Unsöld says that Einstein had run out of new ideas by the age of 45, for which reason he turned his attention to public affairs and in particular to "Jewish nationalism", put on a par with the attachment of other physicists (Lenard, Stark and Wien, for example) to German nationalism. The article says that Einstein's move to Berlin in 1914 was motivated by his "enormous salary demands", and that he was frustrated at not being awarded a Nobel Prize sooner because he needed the prize "in order to divorce his first wife".

The article also says that Einstein

travelled outside Germany with increasing frequency after 1920 "to fill the emptiness created by the drying up of his more profound ideas" and to "make propaganda" for the theory of relativity and for the creation of the state of Israel. Unsöld says that Einstein should have foreseen that his letter to President Roosevelt in 1939, confirming the opinion that nuclear weapons could be built, would lead to the destruction of Hiroshima and Nagasaki. He goes on to urge sympathy for those physicists who stayed behind in Germany after the arrival of Hitler, and who were able to resist the penetration of university faculties by "party personalities" while safeguarding the traditions of research and the university libraries.

There has been a torrent of protests against the publication of Unsöld's article, especially from German physicists working abroad. Unsöld has been charged with antisemitism and with mischievous inaccuracy. In reply to one of these complaints, Professor Rollnik explained that he had not seen the article before publication, that its appearance could be accounted for only by a "gross failure" of coordination between the editorial board of the journal (appointed only in 1980) and the editorial staff and that new rules have been devised to prevent the repetition of "such a nasty event".

The disclaimer now published in the March issue of *Physikalische Blätter* says that the journal must be free to publish controversial opinions even when most members of the society disagree, but that Unsöld's article is inaccurate and that its author has only harmed himself in his attempt to defend physicists who worked in Germany during the Third Reich. It remains to be seen whether it will suffice to bring the row to an end.

on applied biology — to which the Brussels programme is directed — depends on the scale and quality of communications among scientists, and that the principal contribution of the biomolecular engineering programme would have been to set up a network of working contacts throughout Europe. A single European state does not have sufficient biotechnologists to make significant advances in competition with the United States or Japan, they argue.

This point appears to have been lost on the research committee of the Council of Ministers, which has now passed the problem to the diplomatic level in Brussels — the committee of permanent representatives, Coreper. But with the present low level of diplomatic concord in the Community, there seems little hope of a resolution.

● Britain's recent white paper on biotechnology was roundly condemned by British researchers at the second European conference on biotechnology at Eastbourne last week. Dr Duncan Davies, whose chief scientist's office at the Department of Industry produced the white paper, made a flying visit to the meeting to defend his position. He was subjected to a barrage of speeches from the floor, but managed to squeeze in that he would review the white paper's provisions on the education and training of biologists. Informed opinion, however, expects little to come of such a review. **Robert Walgate**

European Space Agency Space on the cheap

Paris

The member states of the European Space Agency are still at odds about the future. Although they have agreed to the proposal of Mr Erik Quistgaard, the agency's director-general, that the budget should be reduced to 60 per cent of its present level over the next ten years, they are hopelessly divided on how to spend the money. The problem is that the reduced budget will not support the three programmes that have been put forward — in telecommunications, space transportation systems and science. So the agency is wrestling with priorities.

Quistgaard's suggestion that the mandatory science budget, to which all member states must contribute, should be increased by 60 per cent over the next ten years has so far fallen on stony ground. Britain, Sweden, Spain and Belgium have flatly turned down the proposal, only Germany has agreed to contribute in full.

Lacking unanimous agreement, the agency may therefore have to take the unprecedented step of admitting optional contributions to scientific programmes. That, according to Ernst Trendelenberg, director of scientific programmes, could divide European space scientists into privileged and under-privileged depending on their nationality.

Germany and France are the most likely to favour this approach, but for different reasons. Germany would like to divert some of the money spent on Spacelab to science, and France is keen that there should be a steady stream of satellites for launching with Ariane, whose development it has supported heavily. Britain, on the other hand, says the agency is too inefficient to justify more spending on science.

The most heated debate in the agency, however, is about the choice of space applications for the next ten years. The two chief candidates — space transportation systems and telecommunications — have been proposed by France and Britain. France argues that Europe must have a launcher to rival the space shuttle, especially if materials processing experiments on Spacelab demonstrate that space has commercial potential. It would like to develop more sophisticated versions of Ariane and ultimately an unmanned partly-reusable space transportation system called Solaris, which is expected to be put before the agency's council at the end of the year.

France disagrees with Britain's proposal to spend more on telecommunications on the grounds that that should be left to industry. Both France and Germany withdrew from early plans to build a large satellite for direct broadcasting, preferring instead a bilateral programme. Neither believes that the "Large Satellite Programme" (L-sat), due to be approved

Quistgaard awaits orders

Paris

Erik Quistgaard, director-general of the European Space Agency, is optimistic. The present stalemate between the member states, he says, might be broken if the users of space technology were more intimately involved. Thus post offices, users of remote sensing data and industrial users of experiments on Spacelab rather than government departments should have more influence over what the agency does. The problem is that although most states know what they want from space, few are decided about their expectations of the agency.

Mr Quistgaard believes that it is up to member states to tell him what they want. Some delegates, however, are looking to him for solutions. The rejection of his proposal that the science budget should be increased is hardly encouraging.

Quistgaard believes that Europe can and should compete with the United States and others in telecommunications, remote sensing and space transportation systems. He favours unmanned space transportation systems as an alternative to the space shuttle. **Judy Redfearn**

in June by Britain, Italy and some of the smaller countries, goes sufficiently beyond their own plans to justify participation. Britain, on the other hand, with a smaller space budget than either France or Germany, sees L-sat as its opportunity to compete for world markets.

The hope is that these two conflicting views can be reconciled. Professor Hubert Curien, director of the Centre National d'Etudes Spatiales, believes that Britain may be moving towards the view that space transportation systems must be developed within the agency. And John Hawkes, of the British Department of Industry, thinks that his French colleagues may now be more ready to agree that telecommunications development beyond L-sat should be supported on a European scale.

If Mr Quistgaard is to put his ten-year plan into operation, he needs to know soon which of these options the agency should be studying. To pursue both would entail scrapping the agreement reached on the future budget. **Judy Redfearn**

Taking note, or shelving?

The British House of Commons last week "took note" of the European proposed programme on biomolecular engineering, a mechanism which allowed the Under Secretary of State for Industry, Mr Michael Marshall, to state the government position on the latest version of the programme.

The original proposal (for 26 million European Units of Account (EUA) over five years) was "over-ambitious in both scope and expenditure", said Mr Marshall, but the shift of emphasis away from research and towards training in the revised proposals (an 11.8 million EUA four-year compromise proposed in Brussels) "is of an extent not hitherto contemplated".

Without naming France, Mr Marshall indicated that the government did not share the French view that there is a significant lack of training opportunities. Nevertheless Britain would continue to work towards an acceptable compromise, he said.

Tam Dalyell MP, described by one member of the House as "almost a walking Select Committee on Science and Technology", warned the minister to be wary of the French position, saying that the new proposals "by pooling training, hand over British expertise to the French for little or no return". The House took note. **Robert Walgate**

Cystic fibrosis

Diagnostic hopes

Techniques now being developed at various centres in Britain and the United States have stimulated optimism about new approaches to the prevention of cystic fibrosis. Workers in the field are particularly excited by the prospect of being able to detect carriers of the genetic defect.

Among genetic deficiency diseases, cystic fibrosis is peculiar in that the nature of the underlying molecular defect is as yet unknown. The belief that only a single gene

is involved has, however, recently been confirmed by Romeo (Bologna), working with records of dispensations for cousin-cousin marriages accumulated at the Vatican. Among Caucasian populations, the frequency of carriers of defective genes is thought to be 1 in 20. The Vatican records have confirmed that cystic fibrosis is an autosomal recessive disease, accounting for the normal incidence in Northern Europe of 1 in 1,600.

Attempts at antenatal diagnosis of cystic fibrosis have been based on the virtual absence, in those with the overt disease, of an uncharacterized proteolytic enzyme. This technique, originally developed by Dr Henry Nadler in Chicago, is now being taken up at other centres. It may eventually be possible to offer amniocentesis to pregnant women who have already produced one child with cystic fibrosis, although the frequency of false positives and false negatives has not yet been determined confidently.

Two techniques for the recognition of carriers are being developed, one of which is based on the occurrence in the serum of both homozygotes and carriers of a protein apparently characteristic of cystic fibrosis. Last year, Dr David Brock's group at the Department of Human Genetics at the University of Edinburgh reported the raising of antibodies against the protein (Manson, Jean and Brock, David J H, *The Lancet*, 16 February 1980) and the possibility that these might be used to distinguish between cystic fibrosis patients (carrying a full dose of the protein), carriers (with a smaller amount) and normal people. Since then, better techniques for raising antibodies have been developed, and there are plans for the development of monoclonal antibodies. The second technique, developed in Los Angeles, is based on the differential response of cystic fibrosis patients, carriers and normal people to the administration of the drug ouabain.

The hope now is that a relatively straightforward way of recognizing carriers may reduce to manageable proportions the provision of amniocentesis for mothers at risk of producing children with cystic fibrosis. In Britain, the Department of Health is under some pressure to prepare for a pilot scheme against the time when the diagnostic techniques are fully developed.

Genetic screening may well be possible before the nature of the underlying defect is understood. One possibility is that the protein characteristic of cystic fibrosis is in normal people removed by the proteolytic enzyme which is the basis of proposed diagnosis of the disease by amniocentesis. In Britain, Professor R. Williamson (St Mary's Hospital Medical School) has embarked on a programme for the identification of the cystic fibrosis gene using techniques of chromosome sorting and genetic manipulation.

Financial support for most of the British research in this field is provided by the

Cystic Fibrosis Research Trust, which is now spending more than £500,000 a year. It is a sore point with those involved that the Medical Research Council's contribution to work in the field, although small, was exaggerated in a reply to a parliamentary question last October.

Safety at Windscale

Yes we were wrong

The British nuclear industry's site at Windscale, the source of a long sequence of mishaps in the past decade, was last week given an adverse report by the Health and Safety Executive, the organization in charge of the Nuclear Installations Inspectorate. And British Nuclear Fuels Limited, the owner and operator of the site at Windscale, put its hand on its heart and acknowledged that it had been at fault.

The report by the Health and Safety Executive is largely concerned with managerial failures at Windscale. It says that British Nuclear Fuels had allowed the condition of the several plants at Windscale to deteriorate until by the early 1970s, their safety could not be assured, and that the efforts made since 1974 to enhance the safety of the Windscale facilities further complicated the assurance of safety by their demand on resources. The report complains that the company paid too little attention to the filling of senior posts at Windscale and that the arrangement whereby the Northwest Area General Manager is responsible for safety in the whole of Windscale but also at other plants means that his responsibilities are too much diluted.

The report says that most incidents at Windscale have arisen because of mistakes in the execution of routine tasks. About a quarter of the 30 or so incidents reported each year since 1976 have involved the exposure of people to doses of radiation exceeding the limits laid down. The inspectors say that the managers of the plant must provide workers on the site with regularly updated and more explicit instructions for carrying out routine tasks and that a promised review of safety procedures should be completed more quickly than now seems likely. In the long run, the Nuclear Installations Inspectorate wants to see a system for the independent audit of safety at Windscale.

The effect of the Health and Safety Executive's report on the publicly owned parts of the British nuclear enterprise is likely to be far-reaching. British nuclear engineers appear to acknowledge that Windscale, because of the way in which it has been created by the addition of new plants to an already overcrowded site, and because of its relative antiquity, is not the ideal model for the safe operation of nuclear plants. But other nuclear sites are likely to encounter similar problems which will be less readily excused.

US university regulations

Complaints listed

Washington

Jumping on the anti-regulation bandwagon which helped propel Mr Ronald Reagan into the White House last November, US colleges and universities are eagerly compiling a list of federal regulations which they feel are stifling the productivity of their educational and research efforts.

One regulation that the universities have in mind is the notorious Circular A-21, issued by the Office of Management and Budget, laying down strict accounting rules which must be obeyed by any scientist and his or her institution receiving federal research support. "A-21 is at the top of our priority list for reform", Mr Sheldon Steinbach, general counsel of the American Council on Education, said last week.

Another rule which the universities would like to see changed is the requirement that scientists should provide a precise account of the way that their time is divided, not only between teaching and research, but also between different research projects. Another relates to cost-sharing between universities. And although the Carter Administration, in its last month in office, took steps through the Office of Management and Budget and the Office of Science and Technology Policy to reduce the impact of the A-21 rules, many universities feel that they are still far from satisfactory.

The universities' point of view had already been presented to Vice-President George Bush, who has been appointed by President Reagan to head a task force on regulatory reform, and has announced the Administration's intention to defer several new regulations proposed by Mr Carter, and also to take a close look at several existing rules to see if they should be revised.

Last week, Mr Bush was presented with a nine-point plan to streamline federal rule-making which had been drafted by 62 leading businessmen and educators organized by the American Council on Education into a Business Higher Education Forum. The forum's statement linked together the complaints made by the business community about, for example, the impact of health and safety regulations, and those coming from the academic community about the general burdens of educational regulation.

The report of the forum says that, as a result of the many educational reforms introduced over the past twenty years, from rules about facilities for handicapped students to affirmative action programmes, regulation of higher education "has mushroomed", with 400 laws on the statute book overseen by 34 committees of Congress.

One complaint is the economic impact of the regulations. According to Howard Bowen, an economics professor at the

Claremont Graduate School, federal regulations alone inflate costs on 3,200 US campuses by an estimated \$3,000 million a year. And the Business Higher Education Forum quotes the widely used — although often challenged — calculation by economist Murray Weidenbaum, now chairman of the Council of Economic Advisors, that federal regulations may be costing US industry \$100,000 million a year overall.

But there are other factors at work as well. "The most important of the regulations' effects is the growing and unwelcome degree to which government intrudes into the private sector, compromising the capacity for independent action in business, labour, academic institutions and elsewhere", says the forum report.

The Bush task force has already said it plans to review two existing regulations covering research. One requires college professors and researchers receiving federal grants or contracts to obtain government approval for leave of absence, changes in research methods or transfer of the research to others, the other requires colleges and universities to share certain percentages of the cost of research financed by the federal government.

Research administrators are cautiously optimistic that, given the general anti-regulatory stance of the new Administration, they will achieve significant changes, a step towards what the American Council on Education forum describes as "a more systematic, rational national policy on regulation".

David Dickson

French nuclear power

Close to the brink?

Brussels

The breakdown in the electricity supply from the French national grid to the nuclear reprocessing plant at Cap de la Hague, at the tip of the Cherbourg peninsula, which occurred in April last year (see *Nature* 24 April 1980, p 653), came within an ace of becoming a major catastrophe. That, at any rate, is what the Brussels-based international socialist group Agenor is claiming.

Agenor has published an account of the incident based on the research of two groups of critical scientists, GSIEN in France, and PERG (Political Ecology Research Group) based in Oxford. The report claims that when the electricity supply was restored the emergency generator "blew", putting all the plant's electrically operated mechanisms out of action. By pure chance one of the worst consequences was avoided. Ten minutes earlier a 30-kg load of plutonium was being prepared. Without the electricity, the agitation needed to prevent the plutonium going critical would have stopped.

The power failure also stopped the cooling of the reservoirs containing radioactive waste from all of France's reactors.

Nuclear Structure Facility

More delays ahead

Britain's Nuclear Structure Facility, under construction at the Science Research Council's Daresbury Laboratory in Cheshire, has run into new problems. The 30-MV tandem Van de Graaf accelerator had been expected to produce its first beam for experiments this month, but the latest setback means that full operation is unlikely before the autumn. The problem is dust, which caused a short circuit across the laddertron — the aluminium and plastic belt which carries charge up the 40-m column to terminals — during a recent trial. The resulting break in a couple of plastic links in the laddertron has been relatively simple to repair, but getting rid of the dust is proving more difficult.

The Nuclear Structure Facility, which should provide a variety of ion beams for nuclear structure studies, has had a long, and not always easy, history. The Science Research Council bit off rather more than it could chew when it decided in 1972 to build a tandem Van de Graaf accelerator that exceeded the capabilities of machines operating at the time. The highest voltage achieved by standard machines was then about 15 MV. Design and technical problems have considerably increased the

Within 10 hours the waste would have begun to boil, releasing caesium-137 and rubidium-106 and making it difficult to gain access to the plant to restore the current and restart the cooling. If that had happened, within three days a radioactive cloud would have formed, and the prevailing southwesterly winds might have carried the cloud as far as London. The uncooled and dry waste left in the reservoir would then have melted down, reaching the water table and causing a steam explosion. At the Cap de la Hague site not only radioactive waste but also stored plutonium would have been distributed over a large area.

On 6 January this year a different accident occurred, the news of which the authorities had more difficulty in suppressing. A fire broke out in a trench used for dumping cladding removed from around the used fuel rods of graphite gas reactors. When alarm bells rang in several workshops on the site, the reaction was to disconnect them. Only later was it realized that radioactivity was coming into the workshops from outside. The fire burned for 15 hours. According to measurements taken by the unions, the level of beta radiation on the site had risen to 26 times the permitted level and alpha radiation to 6 times the permitted level. According to Agenor, the French Atomic Energy Authority (COGEMA) avoided informing the workers downwind of the fire and denied there was any danger to the local population. The radical trade union CFDT had produced figures to show that milk has been contaminated throughout the

building time and the cost.

The dust problem first became apparent last November. There are two sources of dust — beds of alumina pellets used to dry out sulphur hexafluoride gas which surrounds the tandem, and the bearings used in the construction of the machine. A short test last summer revealed no dust problem. For that test, the machine was run without drying the sulphur hexafluoride insulating gas and, according to Dr R Voss at Daresbury, not for long enough for the dust from the bearings to become a problem.

Work is now concentrated on removing all dust from the machine. Dust from the alumina beds should be minimized by passing the gas through more slowly, but it may be more difficult to eliminate dust from the bearings. However, after a few hundred hours of operation, bearing dust should fall off dramatically. After that, the plan is to catch residual dust in traps built into each rung of the laddertron.

The result is that the facility will be late coming on line. It is now hoped that the high-voltage performance can be demonstrated within one or two months, the accelerator tube within the following two months and that commissioning will be complete two months after that. The first experiments are likely in September or October.

Judy Redfearn

Cherbourg peninsula

The intention of GSIEN, PERG and Agenor is to discredit the technology of the Cap de la Hague reprocessing plant and to go on to show that the nuclear power strategy of France as well as Western Europe is wrong.

As well as having problems with safety, the Cap de la Hague plant is not fulfilling its tasks. Cap de la Hague started processing oxide fuel in September 1976, a year behind schedule. The planned capacity is 800 tonnes a year but by 1980 the total amount processed was only 340 tonnes. In the meantime, France has gone on negotiating contracts to reprocess the waste from Japan, Germany, Belgium, Spain, Sweden, the Netherlands and elsewhere. To manage what is estimated to be a commitment to reprocess 6,000 tonnes, France is building two more plants, each with an annual capacity of 800 tonnes. Neither is likely to be in operation by 1988, and meanwhile the waste is being stored.

As well as being behind in coping with nuclear waste, Cap de la Hague is not producing enough fuel. In 1980 the plant handled enough uranium oxide to produce only one tonne of plutonium. Agenor argues that at this speed the fast-breeder strategy falls flat on its face. The initial fuel load of plutonium needed at Malville, a 1,300-MW reactor, is 4.5 tonnes, so that the idea that France or Europe can become self-sufficient with fast-breeders is implausible. Europe boasts only one other reprocessing plant, that at Windscale in the United Kingdom.

Jasper Becker

CORRESPONDENCE

Genetic weaponry?

SIR — The problems of implementing the Helsinki accords in Soviet Russia have been raised once more in your issue of 15 February (p 437) where the case of David Goldfarb was mentioned. I have known David Goldfarb for many years. He was head of the Laboratory of Molecular Genetics of Bacteria and Bacteriophage of the Soviet Academy of Sciences. He has asked to emigrate to Israel and a visa was denied on the grounds that "he had access to classified work".

David Goldfarb was never involved in classified work and did not even have security clearance — as so many scientists in Soviet Russia have. He worked on the genetics of bacteria and bacteriophage. If such research is classified, it means that Soviet Russia intends to use molecular genetics for biological warfare.

So far as is known, the USSR Academy of Sciences is consulted when a scientist asks for an emigration visa. One wonders if the Academy will officially and openly subscribe to the decision of the Soviet authorities according to which molecular genetics is a potential weapon and a military secret.

ANDRÉ LWOFF

Institut Pasteur, Paris, France

NRPB and risk

SIR — It must be admitted that one does not know what happens within an organization without being a part of it. It is equally true that being a part does not automatically make one a reliable critic of it. We have these assertions in mind upon reading Mr S. G. Goss's letter on risks at the National Radiological Protection Board (*Nature* 289, 316, 1980).

We have the highest respect for the scientific integrity of Sir Edward Pochin and the late Dr G. W. Dolphin, and their thoughtful attempts to publish reliable information on radiation risks.

We further believe that this view is shared by other US colleagues who have had the opportunity of interacting with these two. Derogatory evidence would have to have a decidedly firmer foundation than that offered in the Goss letter to change this viewpoint.

WILLIAM J. BAIR
ROY C. THOMPSON
HERBERT M. PARKER

Battelle Pacific Northwest Laboratories,
Richland, Washington, USA

Conservation sites

SIR — In his letter of 12 March 1981, Dr C. Muir criticises actions and advice of the Nature Conservancy Council (NCC) over Sites of Special Scientific Interest (SSSI) on the grounds that such sites are misguidedly chosen to protect endangered species, because species extinction is a natural and inevitable process, and that the site notification procedure imposes an unacceptable financial burden on landowners. The NCC does not notify SSSIs from "caprice" but because Parliament in its wisdom conferred this duty upon us, as a key part of a national strategy for nature conservation. We continue to do this because a very large number of concerned people believe that SSSIs are important to the conservation of wildlife and physical features.

Most extinctions during the last 2,000 years have been man-induced, and it is this accelerated loss that we are trying to stem. By

2000 AD the world will have lost no fewer than one million of the plant and animal species existing today, mainly through habitat destruction and over-exploitation. In Britain many indigenous species could become extinct. Whether this is held to be evolution or not is irrelevant: the fact is that many people do not wish such losses to occur and look to NCC and other bodies to prevent them.

SSSIs are in any case not chosen just for the presence of rare or endangered species, but to give a national network of important semi-natural habitats. Many are selected quite separately for the importance of their geological and physiographic features. The term "scientific interest" is interpreted broadly as the value of the biological and physical attributes of a site to that informed sector of society which has a concern for such phenomena. The NCC has defined criteria for evaluation of these attributes based on the full range of public interest involved, which is clearly wider than scientists alone.

The SSSI device is one of the few means available for defending the important areas for nature conservation but its actual inadequacy in this respect is currently the subject of lengthy parliamentary debate on the Wildlife and Countryside Bill. Our concern is indeed to safeguard these sites in perpetuity. It is a staggeringly narrow and mechanistic view which holds that once "objects or phenomena" have been "studied" they are of no further value to science and can, by implication, be allowed to disappear. And beyond this, who has the confidence to assert that all possible knowledge has been gleaned from any object or phenomenon?

The responsibility of government to cater for nature conservation as a public interest is not, in a democratic society, altered by the fact that some people find the resulting measures inconvenient to their material interests or irrelevant to their personal beliefs. And just to get things in proportion, the sense of official priority accorded to nature conservation in public affairs is reflected in the £10 million annual budget of NCC, compared with £700 million allocated to agricultural support and £9,500 million to defence.

The NCC has long accepted that when owners and occupiers of land are asked to incur a material loss in order to sustain a nature conservation interest, the principle of compensatory payment is entirely fair. This principle would remain an integral part of the additional site safeguard measures which NCC has been seeking in the Wildlife and Countryside Bill. It seems equally reasonable to ask that owners and occupiers of land accept that they have, on behalf of the nation, a responsibility in stewardship for the national heritage of nature which is now suffering such severe and accelerating losses. We are all increasingly enmeshed in a single society, and when the owners and managers of our national resources ask the taxpayer to help their private circumstances by direct or indirect cash hand-outs, another point of principle arises. This is that these parties should also accept that such funds are distributed across the whole wide spectrum of taxpayer interests, according to a balance of adjudication through the proper agencies of government, of which NCC is one.

D. A. RADCLIFFE

Nature Conservancy Council,
London SW1, UK

Selling Darwin

SIR — So much has been written on Darwin and Marx, since Halstead's original diatribe against the use of cladistics in the displays of the Natural History Museum (*Nature* 288, 208, 1980), that I wonder whether these two venerable gentlemen are not now reaching for their pens to set the record straight. Until *Nature's* Olympian warnings (26 February, p 735) on the "selling out on Darwinism" in the museum (and by association other new evolutionary approaches) one felt safe from the bogies in Halstead's head. Now that they are out and about rather than dead and buried, it is amusing to consider the original error which gave rise to them.

Halstead's polemics "follow like a tedious argument of insidious intent" (T. S. Eliot). His accusation that cladists and proponents of "punctuated equilibria" are furthering the Marxist concept of dialectical materialism reveals his ignorance of the fundamentals of this concept and of Darwinism. If it is necessary to find a biological phenomenon in keeping with dialectical materialism, then Darwin and not "punctuated equilibria" provides a perfect example. According to dialectical materialism (unsullied version) there are two inter-related components to every process in that a gradual accumulation of small quantitative changes leads inevitably to a qualitative change of state. Engel's favourite scientific example was the boiling of water. A qualitative change from liquid to gas only occurs after a sufficient quantitative increase in the thermal agitation of the molecules has taken place. According to popular Darwinism a gradual quantitative accumulation of allelic differences as an adaptive response to a specific environment leads inevitably to a qualitative change of state (the inception of reproductive isolation/speciation). If dialectical materialism upsets Halstead he would do better to dismantle 130 years of Darwinism.

Ironically the precise mechanism by which natural selection could increase reproductive isolation perplexed Darwin and, despite the subsequent attempts of Fisher, Muller and Dobzhansky to explain how this might happen, we have little empirical data. "Let me first say that no man could have more earnestly wished for the success of Natural Selection in regard to sterility than I did, I always felt sure it would be worked out but always failed in detail" (Darwin letters to Wallace and Huxley). Notwithstanding these doubts and lack of data Darwinism flourishes today as an explanatory mechanism of adaptation that leads to speciation.

The theory of "punctuated equilibria" does not see species formation necessarily as the inevitable consequence of microevolutionary adaptive differences. It dissociates (to an extent) quantitative (microevolutionary) from qualitative (macroevolutionary) phenomena. Thus with regard to the process of adaptation, it is Darwinian. With regard to the process of speciation, there are several current notions (ranging from possible developmental constraints on phylogenetic progress to possible biological effects of saltatory and rapidly spreading changes in genome organization) that, although novel to a degree, leave no room for instant creationism or the demise of Darwin, in the way it is construed.

They are basically concerned with differential rates of evolutionary change, and the precise role of selection has not, as yet been ascertained. We should not be too surprised if speciation turns out to have an accidental as well as an adaptive component. The imaginative and enquiring spirit of Darwin lives in all except the minds of those with other axes to grind.

I have a feeling that Darwin would be overjoyed at the general acceptability of his concept of fine-grained adaptation under natural selection, and would not be unreceptive to its possible dissociation as a process from that of speciation. Marx, on the other hand, could but grunt that a dissociation of the two processes no longer requires the inevitability of a change of state of one species into another.

GABRIEL DOVER
Department of Genetics,
University of Cambridge, UK

From the museum...

SIR — As music is the food of love so surely debate is the nourishment of science, but the crashing cymbals of your leading article of 26 February were a discordant accompaniment to a meagre meal¹. Your subsequent article² sets the stage for a more reasoned discussion.

Controversy and debate over the hypotheses associated with the concept of organic evolution have recently received considerable attention. It can hardly be described as a rot permeating just the British Museum (Natural History) but rather the open questioning of evolutionary theory by serious biologists and philosophers — a fact apparently denied in your first leading article. Furthermore there are clear indications of unease amongst educationalists^{3,4} at the way in which the concept of evolution is being taught, not as a hypothesis open to criticism but as a fact.

Many practising biologists and teachers echo Olson's remark "Evolutionary theory should be treated like any other scientific theory, as a matter for dispassionate and objective study of the evidence available"⁵. Yet this only reflects the more pungent and provocative comment made by Darwin's "bulldog" over a century ago "the scientific spirit is of more value than its products, and irrationally held truths may be more harmful than reasoned errors. Now the essence of the scientific spirit is criticism"⁶.

The views of many scientists in this museum are encapsulated in a quotation from the museum's "recent brochure"

Biologists try to reconstruct the course of evolution from the characteristics of living animals and plants and from fossils, which give a time-scale to the story. If the theory of evolution is true, the features used to classify species in groups were acquired by the common ancestor of the group and inherited by the living descendants. In this light, the groups-within-groups of classification are seen as the descendants of more or less remote common ancestors. And classifying animals and plants is a way of expressing ideas about the course of evolution. This is why classification is interesting and important to biologists, and why the work of classifying is never finished. For new discoveries lead to new ideas about the course of evolution.

Our understanding of the theory of evolution is that it is an amalgam of many subsidiary hypotheses, related not only to the patterns of

organic diversity but also to the mechanisms whereby they arise. Here we are faced with a core theory and its protective satellite belts. Clearly various aspects of these hypotheses have been criticized and improved upon, and there is no reason to believe that this evolution has stopped, nor should we pretend that it has stopped. As practising systematists we recognize and we do not underestimate the enormous heuristic value of the theory as a stimulus to research.

The theory of evolution must surely be considered an "open question" and we therefore agree that "in the public presentation of science, it is proper whenever appropriate to say that disputed matters are in doubt"⁷. We cannot see, however, how the views expressed in the museum's brochure indicate any divergence from such an objective. Rather, they make a positive statement about the utility of the theory as an explanation for the diversity of life.

P H GREENWOOD
J F PEAKE
Department of Zoology,
British Museum (Natural History)
London SW7, UK

- 1 *Nature* 289, 735 (1981)
- 2 *Nature* 290, 75–76 (1981)
- 3 Hutchinson, C. *Biologist* 27, 252–256 (1980)
- 4 Harper, G. H. *School Science Review* 87, 258–268 (1977)
- 5 Olsen, E. C. *The Evolution of Life* (Weidenfeld and Nicolson, London, 1965)
- 6 Huxley, T. S. *The Coming of Age of the "Origin of Species"* (reprinted in *Darwiniana*, Macmillan, London, 1907)

SIR — Dr Colin Patterson¹ says that cladistics is "not about evolution". Cladistics may therefore be considered essentially a method of classification. Based on morphological characters it arranges taxa into a hierarchy and the hierarchy itself has no necessary phylogenetic implications². Since cladistic classifications are based on "derived" similarity, it seems to have been accepted³ that they may resolve the classification hierarchy more fully than traditional classifications, which are based on "clustering" similarity⁴.

If this is all there is to cladistics then at least its taxonomic substance seems uncontroversial. However, any such hierarchical classification, whilst practically useful, is essentially non-scientific⁵ as the aim of science is to explain simple taxonomic relations, then to test the explanation (scientific theories) to see if they make sense⁶. Evolutionary theory is one such scientific theory⁷.

Cladistic method can be incorporated into evolutionary theory by simply accepting that the cladistic hierarchy represents an assumed time axis, a view taken by Hennigian cladists. A historical sequence is "predicted" from the morphology of organisms and in context of Popperian concepts of science, cladists have gone on to compare the congruence of such predicted sequences with other such sequences based on different, semi-independent lines of evidence, such as historical geography⁸.

However, stratigraphy can also be used constructively in such comparisons to call into question the sequence of taxa presented by cladograms and thus to test the usefulness of the hypothesis outlined above⁹.

Cladists to date have applied the concept of testing mainly from an internal perspective of character analysis¹⁰, and historical time-space coordinates of taxa have in general been ignored. The cladogram has thus been regarded as absolute, whilst evidence from stratigraphy, because it is incomplete, has been recognized as relative — and apparently therefore as worthless¹¹. Little attempt has

been made to hold the cladistic system itself up to test. It seems to have been overlooked that cladistic classification may make morphological evidence seem convincingly complete, even though it is not actually complete. I do not understand how a system that is relative (cladogram) can be assumed to be absolute and thus used universally to ignore or "refute" historical evidence that is actually true¹³. The view promulgated here regards both sets of evidence as relative, and thus as commensurable. Cladograms present a relative sequence and stratigraphy presents a relative sequence, and both have their own semi-independent corroborators and falsifiers. In actual practice the testing role of stratigraphy would therefore be difficult, a sign of a mature science^{9,12}, in some cases the stratigraphy may be trusted more than the cladogram, in others the cladogram may be relatively well corroborated¹³.

It is, on the other hand, naively easy to question the cladistic system as a basis for delimiting species. Since cladistic hierarchies reflect characters of organisms, and since such characters are expressed as a hierarchy of proper sub-classes (monothetic sets)⁴, a logical consequence is that classes cannot be discriminated from individuals within the context of character analysis. Men and women form cladistically distinct taxa, and pregnant women a taxon nested within women as a whole. Straightforward genealogical observation refutes this, since a pregnant lady can deliver a son and return to her generalized condition. This argument may seem facile but for the fact that one cladist has recently described several new species out of what many evolutionists might take to be sub-specific populations¹⁴. At any rate there is a potential danger of over-splitting taxa.

My view differs from some of my colleagues in that I maintain cladistics is about evolution and that any sensibly testable system will recognize that evolution is a complex, mature theory, requiring complex modes of testing beyond "naïve falsificationism". Basically I consider evolution and the fossil record as very much the proper context for cladistics and not the converse. It is surely commonsense to use historical evidence to examine history and not to view it askance through a haze of global aristotelian philosophy of character analysis, as transformed cladists appear now to be doing. Transformed cladists may justly have rediscovered Aristotle, but if they decide to wipe out Galileo they reject not only Darwin but all of modern science¹².

CHRISTOPHER HILL
Department of Palaeontology,
British Museum (Natural History),
London SW7, UK

- 1 Patterson, C. *Nature* 288, 430 (1980)
- 2 Platnick, N. I. & Cameron, H. D. *Syst. Zool.* 26, 380–385 (1977)
- 3 Mayr, E. Z. *Zool. Syst. Evol. Forsch.* 12, 94–128 (1974)
- 4 Hill, C. R. & Crane, P. R. in *Problems of Phylogenetic Reconstruction* (eds Joysey, K. A. & Friday, A. E.) (Academic, in the press)
- 5 Hull, D. L. *Evolution* 21, 174–189 (1967); *Syst. Zool.* 17, 438–457 (1968); *A. Rev. Ecol. Syst.* 1, 19–54 (1970)
- 6 Popper, K. R. *The Logic of Scientific Discovery* (Hutchinson, London, 1959)
- 7 Popper, K. R. *New Scientist* 87, 611 (1980)
- 8 Nelson, G. J. & Platnick, N. I. *Bioscience* 30, 339–343 (1980)
- 9 Lakatos, I. *Proc. Aristotelian Soc.* 69, 149–186 (1968)
- 10 Gaffney, E. S. in *Phylogenetic Analysis and Paleontology* (eds Cracraft, J. & Eldredge, N.) 79–111 (Columbia University Press, New York)
- 11 Nelson, G. J. *Syst. Zool.* 27, 324–345 (1978)
- 12 Feyerabend, P. *Against Method* (Verso, London, 1975)
- 13 Manuscript submitted to *Science*
- 14 Rosen, D. E. *Bull. Amer. Mus. Nat. Hist.* 162, 275–278 (1979)

NEWS AND VIEWS

Tissue-specific expression of α -amylase

from R A Flavell

TISSUE-SPECIFIC expression of eukaryotic genes is always intriguing but of particular interest are cases where the expression of a given type of protein is confined to a few tissues, yet where the level of expression in each of these tissues differs. The α -amylase genes belong to such a system. α -Amylase is produced at high levels in salivary glands and the pancreas, but at much lower levels in the liver. Previous studies on the α -amylases of the laboratory mouse have suggested two structural genes, designated *Amy 1* and *Amy 2*. Several recent papers from a group at the Swiss Institute for Experimental Cancer Research describe the molecular organization of these genes. A single gene, *Amy 1A*, encodes both the liver and salivary gland amylases (which are identical proteins). Differential expression of these two mRNAs probably requires the use of at least two different promoters for a single cellular gene. Equally interesting, differential splicing is required for the generation of these two overlapping, but non-identical mRNAs. The second gene (*Amy 2*) encodes pancreatic α -amylase, which has a different primary sequence from the salivary gland/liver forms (Schibler *et al* *J molec Biol* **142**, 93, 1980; Hagenbuchle *et al* *Cell* **21**, 179, 1980).

The notion that differential splicing can be used to regulate gene expression is not new. Indeed, this principle was borne out by the detailed study of viral transcripts where, in several cases, overlapping genes have been found. A well known example is that of the gene products of the early region of simian virus 40, the tumour antigens T and t. t is encoded in a DNA segment colinear with the T gene. Although t mRNA is spliced, this splice occurs after the stop codon and does not, therefore, affect the t antigen protein-coding regions. T antigen is produced from essentially the same mRNA by splicing the region encoding the N-terminal segment of t antigen to the same splice-acceptor sequence as is used in t mRNA. The result is the elimination of the t-antigen stop codon and the fusion, in a compatible reading frame, of two protein-coding segments,

translation of this RNA generates the mosaic T antigen. This is a simple example. The characterization of adenovirus RNAs by a large number of groups has revealed a dazzling spectrum of gene products generated by both differential use of promoters and differential splicing (see review by Ziff *Nature* **287**, 491, 1980).

Viruses pay a premium to keep their genomes small, and overlapping genes are an excellent way to compress a lot of information into a little space (as in the smaller bacteriophages such as ϕ X174). There is little indication that higher organisms have gone to much trouble to trim down the size of their genomes. The average *Nature* reader has 3×10^9 base pairs of DNA per genome, give or take the odd supernumary chromosome, while salamanders, whose aspirations are significantly more modest, possess up to 10^{10} base pairs of DNA per haploid genome. Since it is believed that only about one-tenth of our DNA is used to encode proteins, one might predict that overlapping genes would be restricted to the viruses and that the host genomes would use the more familiar path of gene duplication to generate the same type of diversity. Not so, says α -amylase.

The story runs approximately as follows, with certain details omitted for simplicity.

The primary structures of murine liver and salivary gland α -amylase mRNAs have been compared by sequence and S_1 nuclease/hybridization analyses, and this has been described in a recent issue of *Nature* (Hagenbuchle *et al* *Nature* **289**, 643, 1981). These mRNAs are essentially identical except for the 5'-untranslated region, which is 206 nucleotides long in the case of the major liver species, but only 95 nucleotides in the salivary gland mRNA. Comparison of these sequences shows that both liver and salivary gland mRNAs share the 48 nucleotides of the 5'-untranslated region immediately adjacent to the coding region.

The secret of how this mosaic structure is

generated was revealed by a structural analysis of the chromosomal α -amylase genes. Young, Hagenbuchle and Schibler (*Cell* **23**, 451, 1981) cloned the α -amylase genes, located the mRNA-coding segments of their clones and determined the DNA sequence of the 5' areas of the gene, the region described in the current study covers approximately the 5' one-fifth of the mRNA-coding region. It turns out that the liver and salivary gland α -amylases are encoded by the same mosaic gene. Four exons have been found so far (there are presumably several other exons in the 3' regions of the gene, but these have not been analysed in mouse, for the rat, see MacDonald *et al* *Nature* **287**, 117, 1980) which correspond, respectively from 3' to 5', to an internal coding region (called exon 4 here), the 5'-terminal coding region plus the 45 nucleotides of the 5'-untranslated region (exon 3), common to both mRNAs, and two alternative 5'-terminal exons — one of 161 nucleotides (exon 2), which is utilized in liver α -amylase mRNA, and one, about 2.8 kilobases further to the 5' side in chromosomal DNA, of 50 nucleotides (exon 1) which is used for the salivary gland mRNA. The two alternative mRNAs are probably obtained by differential transcription and/or processing but these points remain to be elucidated. There are promoter-like TATAA boxes at the usual 24 (± 1) nucleotides upstream from both exons 1 and 2 — it is plausible⁵ (see Young *et al op cit*), therefore, that transcription initiates at exon 1 in the salivary gland and exon 2 in the liver. Equally, transcription could initiate at both sites and only the mRNA containing the appropriate leader might be exported to the cytoplasm in the relevant tissue. Whatever the mechanism, the uncapped liver-type leader sequence (exon 2) must presumably act as an intron in the salivary gland transcript, differential splicing must ultimately generate tissue-specific mRNAs with exons 1, 3 and 4 being spliced in the salivary gland and 2, 3 and 4 being spliced in the liver. As Young and his colleagues point out, this type of situation also exists in the immunoglobulin genes where multiple J regions exist in the active,

R A Flavell is at the National Institute for Medical Research, London

rearranged, K-chain gene, for example, only one of these is destined to appear in the mRNA, the remainder being removed by differential splicing

Why does the mouse go to all this trouble? Any explanation should consider the fact that α -amylase mRNA accounts for 2 per cent of the cytoplasmic mRNA in the salivary gland, but only 0.02 per cent in liver. This level may reflect the transcription rate from the two genes, the 'salivary gland' promoter would then be considerably stronger than the 'liver' promoter. Alternatively, the rate of RNA processing and/or export to the cytoplasm may be different for the two mRNAs. Another clue may reside in the fact that the liver mRNA has a much longer 5'-untranslated

region than the salivary gland mRNA. It is possible that this affects the translation efficiencies of the mRNA (see 'Discussion' in Young *et al.*). Although the explanation for this fascinating genetic mechanism is the subject of future work, it is, of course, likely to be tied up with the primary question — what determines the different level of α -amylase in two different tissues?

In the rat (and other mammals) the situation may be even more complicated. MacDonald and his co-workers (*Nature* 287, 17, 1980) have analysed the rat α -amylase genes and these studies point to at least five non-allelic α -amylase genes or pseudogenes. The further elucidation of the expression of the α -amylase gene family remains a mouth-watering prospect. □

Antibody-mediated enhancement of rabies virus

from J S Porterfield

THE long incubation period following naturally acquired rabies virus infection allows time for manipulation of the host's immune system in ways that can promote protection. Louis Pasteur's treatment of Joseph Meister was dramatically successful almost a century ago, and since then, very many others have benefited from rabies vaccines¹. There are, however, occasional failures, even when vaccine of established potency has been used, and much remains obscure about the pathogenesis of rabies virus infection, and about the precise way in which rabies vaccine confers protection. One puzzling effect, seen in both monkeys and mice immunized with rabies vaccine and subsequently challenged with rabies virus, has been termed the 'early death' phenomenon, since some vaccinated animals die sooner than non-vaccinated controls given the same challenge^{2,3}. Following earlier work which indicated that both bone marrow-derived B lymphocytes and thymus-derived T lymphocytes contributed to the clearance of rabies virus from the central nervous system^{4,5}, Prabhakar and Nathanson now present evidence, on p. 590 of this issue of *Nature*, that 'early death' in mice is mediated through antibodies, and is not a cell-mediated effect involving T lymphocytes, they do not, however, explain the mechanism whereby antibody produces this paradoxical response.

Apart from rabies, there are a number of other infections in which antibodies appear to produce effects which are detrimental to the host. The increased risk of severe bronchiolitis and pneumonia seen in infants who had been vaccinated with killed respiratory syncytial virus vaccine

and subsequently acquired natural infection with that virus is well documented, but still inadequately explained⁶. Dengue viruses normally produce a non-fatal illness with a rash, but a proportion of children in South-East Asia develop haemorrhagic manifestations and may die from shock⁷. Although the effector mechanism responsible for the shock remains controversial, several laboratory studies related to haemorrhagic dengue are relevant to the 'early death' phenomenon. Monkeys inoculated with dengue virus mixed with sub-neutralizing concentrations of anti-dengue antibodies show higher levels of viraemia than control animals given dengue virus only⁸. Dengue virus replicates poorly in normal human or monkey peripheral blood white cell preparations, but shows enhanced yields of virus when exposed to infectious virus-antibody mixtures⁹. This antibody-mediated enhancement of viral replication is not restricted to dengue virus, but has also been demonstrated in continuous cell lines of mouse and human macrophages with West Nile virus¹⁰, with other Togaviridae and Bunyaviridae¹¹. Halstead first suggested that the mechanism involved in the human monocyte system depended on an increased uptake of virus complexed with antibody through Fc receptors (present on these cells but absent from most other cells), the internalized virus escaping destruction by the phagocytes and replicating within them⁹. Using the simpler model system with mouse macrophage-like cell lines, Peiris *et al.*¹¹ showed that the antibody-mediated enhancement of West Nile virus replication was specifically blocked by monoclonal antibodies directed against Fc receptors on the cells they were using.

Are these findings at all relevant to the rabies 'early death' effect? Rabies viruses were thought to be serologically identical, but a number of rabies-related viruses are now known¹², and antigenic variation between rabies virus strains has been established¹³. Mice inoculated with Lagos Bat or Mokola viruses, two of the rabies-related viruses, and subsequently challenged with rabies virus, also died more quickly¹⁴. These findings suggest that it is not essential to have homologous neutralizing antibodies to produce the rabies 'early death' effect, but that cross-reacting sera may also be active in this system as in the dengue system^{10,15}. Three monoclonal antibodies directed against the envelope glycoprotein of West Nile virus will all enhance West Nile virus although only one of these has neutralizing activity¹⁶. Monoclonal antibodies against known components of rabies virus and against rabies-related viruses are available, and it would be of great interest to determine which of these was capable of producing the 'early death' effect with rabies virus.

It might be objected that rabies virus replicates in the central nervous system in cells that would not have Fc receptors, so that the mechanism of opsonization of virus-antibody complexes could not apply. Rabies virus is capable of replicating in a wide variety of different cell types in addition to neurones, and it is certainly possible that macrophages within the brain or spinal cord might support replication, following 'enhanced' uptake in the presence of anti-rabies antibodies, although experiments with mouse peritoneal macrophages lend little support to this view¹⁷.

These comments in no way detract from the unquestionable efficacy of present post-exposure treatments of rabies, which combine the use of human diploid cell vaccine with human rabies immune globulin¹⁸⁻²⁰. They emphasize a continuing need for investigation of the complex interactions involved in the pathogenesis of rabies virus infection. □

- 1 Pasteur, L. *C r hebdo Séanc Acad Sci, Paris* 101, 765 (1885).
- 2 Sikes, R K *et al Bull Wild Hlth Org* 45, 1 (1971).
- 3 Blancou, J, Andral, B & Andral, L *J gen Virol* 50, 433 (1980).
- 4 Miller, A, Morse, H C III, Winkelsstein, J & Nathanson, J *J Immun* 121, 321, (1978).
- 5 Smith, J S *Infect Immun* 31, 297 (1981).
- 6 McIntosh, K & Fishaut, J M *Prog med Virol* 26, 94 (1980).
- 7 Halstead, S B *Bull Wild Hlth Org* 58, 1 (1980).
- 8 Halstead, S B *J infect Dis* 140, 527 (1980).
- 9 Halstead, S B & O'Rourke, E J *J exp Med* 146, 201 (1977).
- 10 Peiris, J S M & Porterfield, J S *Nature* 282, 509 (1979).
- 11 Peiris, J S M, Gordon, S, Unkles, J C & Porterfield, J S *Nature* 289, 189 (1981).
- 12 Matthews, R E F *Intervirology* 12, 132 (1979).
- 13 Wiktor, T J & Koprowski, H *J exp Med* 152, 99 (1980).
- 14 Tignor, G H, Shope, R E, Gershon, R K & Waksman, B H *J Immun* 112, 260 (1974).
- 15 Flamand, A, Wiktor, T J & Koprowski, H *J gen Virol* 48, 105 (1980).
- 16 Peiris, J S M thesis, University of Oxford (1980).
- 17 Turner, G S & Ballard, R *J gen Virol* 30, 223 (1976).
- 18 World Health Organization Working Group 2 *Dev Biol Stand* 40, 268 (1978).
- 19 Aoki, F Y, Tyrrell, D A J, Hill, L E & Turner, G S *Lancet* i, 660 (1975).
- 20 Anderson, L J *et al J infect Dis* 142, 133 (1980).

J S Porterfield is in the Sir William Dunn School of Pathology, University of Oxford

Antibodies, introns and biosynthetic versatility

from Miranda Robertson and Mike Hobart

SOON after the discovery that eukaryotic genes may be interrupted by non-coding DNA sequences, Gilbert¹ speculated in *Nature* that this unexpected arrangement might make it possible to produce variants of a single protein by differential splicing of the interrupted RNA. Specifically, he suggested that such a mechanism might underlie the well defined sequential changes that take place in the immunoglobulin molecule in the course of B lymphocyte ontogeny. As a result of these changes, a single antigen-binding site is associated successively with different effector sites to produce first immunoglobulin M, then immunoglobulin D, and later immunoglobulin G or sometimes A or E. Each of the different classes of immunoglobulin is expressed first as a membrane-bound receptor molecule and later, on the terminal differentiation of the lymphocyte, as a secreted protein which is identical to the receptor in all but a few residues at its carboxy terminal.

Both the class of the immunoglobulin molecule and whether it is membrane-bound or secreted depend on the structure of the constant region of its heavy chains (see Fig 1). Thus each heavy-chain variable region (V_H region) is associated in turn with a constant region (C_H region) coding for the μ , δ and γ , α or ϵ chains corresponding to IgM, IgD, IgG, IgA and IgE, and each of these chains may then undergo modification at its 3' end. It is now known that the class switch from IgM to IgG, IgA or IgE is mediated by translocation of the V_H region DNA with deletion of the C_H regions 5' to the one that is expressed². Deletion, however, cannot explain the switch from IgM to IgD, because some B cells, including a few tumour cell lines, synthesize IgM and IgD at the same time; nor is it an attractive explanation for the switch from membrane to secreted proteins,

Miranda Robertson is an associate editor of *Nature* and Mike Hobart is at the Medical Research Council Centre, Cambridge

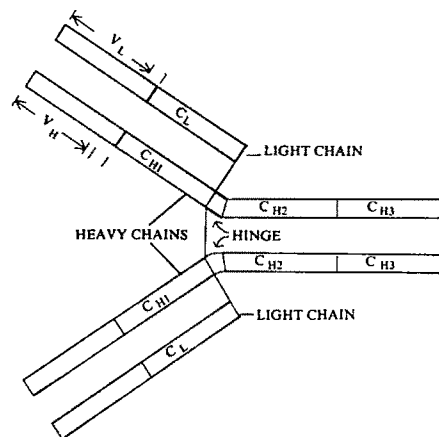


Fig 1 Diagram of an antibody molecule showing the division of the heavy chains into the variable regions that determine antigen-binding specificity and the constant region domains. Each of the four domains, C_{H1} , C_{H2} , C_{H3} and the hinge, is encoded in a separate intron on the DNA. The sequences of the C_H introns are different for the different classes of immunoglobulin. The secreted form of each class (C_s) differs from the membrane form (C_m) in the last few carboxy terminal residues of C_{H3} ; the membrane form has a longer and more hydrophobic carboxy-terminal which anchors it in the membrane, the diagram shows the secreted version.

which are also simultaneously expressed. Instead, as Gilbert surmised, these processes are controlled by processing of the RNA.

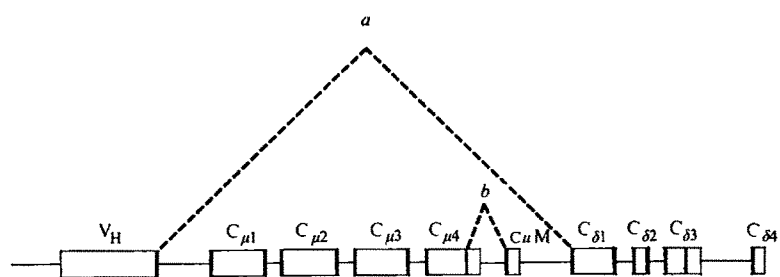
Immunoglobulin D biosynthesis has become accessible to investigation through the work of Liu and his colleagues³ who have located the gene coding for the corresponding δ constant region. The gene itself is very peculiar, having only two constant region domains (instead of the usual three) and an abnormal hinge region. Furthermore, it lies only about 2,000 base pairs from the μ gene corresponding to the IgM constant region; this is a much shorter distance than has so far been found to separate other constant region genes, and it means a single stable transcript could be

made from the combined μ and δ coding regions. There is, indeed, now preliminary evidence for the existence of such a transcript⁴. The switch from μ to δ could thus be mediated by the removal of the entire μ gene as an intron (see Fig 2). The simultaneous expression of μ would be due to simultaneous production of shorter transcripts, ending before the δ sequences, or to polyadenylation at a site between the μ and δ genes.

This scheme has more recently been strongly reinforced by Moore *et al.*, who have confirmed the location of the δ gene in DNA both from mouse sperm and from a mouse hybridoma expressing membrane-bound forms of IgM and IgD, and established that the μ gene is not rearranged in the hybridoma DNA⁵. At the same time, however, they have raised the possibility that plasma cells committed to the secretion of IgD, unlike B cells transiently synthesizing the membrane form, may delete the μ sequences. So far, the evidence rests on a single rat plasmacytoma which secretes IgD and seems to have deleted the μ gene. Whether the deletion of μ is really a normal concomitant of the terminal differentiation of IgD-secreting lymphocytes, or whether it simply reflects the notorious instability of constant region genes in cultured cells, it is too early to say. It is not yet known, for example, whether the specialized switch sequences characteristic of the regions on the 5' side of the μ and γ genes² are to be found between the mouse μ and δ genes, but Rabbitts *et al.*⁶ have sequenced most of the region between the human μ and δ genes without discovering any.

The deletion or otherwise of the μ gene has, however, no direct bearing on the switch from membrane to secreted IgD, which is presumed to occur by the mechanism recently established for IgM. Independent observations by Alt *et al.*⁷ and Rogers *et al.*⁸ show that cells which are making both membrane and secreted IgM

Fig 2 Diagram of chromosomal DNA coding for IgM and IgD in a B cell. The open rectangles are exons and the intervening solid lines represent introns. Heavy vertical bars represent RNA splicing signals. In cells, expressing both membrane IgM and membrane IgD, it is believed that the entire region is transcribed. Dotted line *a* joins the introns that are spliced to produce membrane IgD, with the elimination of all the C_μ introns. Dotted line *b* joins the introns that are spliced to produce the membrane form of IgM. Note that if the switch from membrane to secreted IgM occurs through termination of transcription after the μ_s segment, transcription of the δ gene will also be shut off. The switch from membrane to secreted δ chains is believed to occur as described for μ .



contain μ chain messenger RNAs of different lengths. By examining cells in which the two forms of μ chain are synthesized in different proportions, Alt *et al.* were able to deduce that the longer of the two messages codes for the membrane (μ_m) form. Rogers *et al.* hybridized the two messages with cloned genomic DNA containing the μ gene, and found that while both hybridize the coding sequence of the four major μ domains, the longer one also hybridizes with a short segment on the 3' side. The same team went on to show⁹ that this short segment codes for a hydrophobic protein sequence whose characteristics are similar to those of known membrane-bound proteins.

Thus the secreted form of the protein (μ_s) is encoded by the four main μ gene segments, while the expression of the membrane form entails the addition of a further segment at the 3' side. This 3' μ_m segment has a sequence characteristic of a splice site at the beginning of an exon at its 5' end, but the matching splice sequence upstream occurs within the coding sequence of the C μ 4 domain, followed by the coding sequence for the last 19 amino acids of the μ_s chain and a termination codon. Early *et al.* suggest on the basis of this organization that μ_m , the first antibody to be synthesized by a B cell, is produced from a long transcript extending to a polyadenylation site beyond the μ_m coding segment. In the course of the processing of the transcript to mRNA the μ_s coding segment is eliminated as an intron, with the consequence that the termination codon is also removed.

When the B cell matures into an antibody-secreting plasma cell, either transcription terminates upstream of the μ_m segment or, for some reason, the μ_m segment is cleaved from the transcript. In either case the μ_s sequence is retained in the mRNA and read until the termination codon.

Thus the class switch from IgM to IgD seems to work by a mechanism analogous to that which operates the switch from μ_m to μ_s , and it has, incidentally, been pointed out that the μ_m - μ_s switch may automatically eliminate the production of IgD by IgM-secreting cells if transcription ends at the μ_s segment (see Fig 2).

The organization and regulation of the δ coding sequences open two further questions, the first specifically about the δ chain and the second generally about RNA splicing.

The first arises from the fact that the

coding sequences as described by Liu *et al.* can account for the shorter of the two known versions of the δ chain but not for the longer version¹⁰. However, further sequences apparently coding for the δ chain have more recently been detected (but not yet characterized)¹¹ further downstream, lending weight to the proposal of Liu *et al.* that differential splicing may provide a family of different δ chains with different — and it must be confessed entirely unknown — functions.

The second question is that of the mechanism by which introns are eliminated

from nuclear RNA. Until now, the simplest scheme — that the splice site at the end of each exon is paired with the matching sequence at the beginning of the next one downstream — has also been consistent with the known data. However, it plainly cannot account for the removal of the μ exons from the large μ - δ transcript. There is still a formal possibility that such a combined transcript does not exist, but if it does, it will clearly represent a further strong incentive to molecular geneticists of immunoglobulin to devise a splicing system that will work *in vitro*. □

Anomalous sputtering at high energy

from P K Haff and L E Seiberling

WHENEVER a sufficiently energetic stream of particles impinges on a target, there is some probability that target atoms and molecules will be knocked loose from the surface. This process of erosion on an atomic scale is called sputtering. For incident ions in the keV per AMU range, the average number of target particles ejected per ion is called the sputtering yield, S . Values of S in the range 1 to 10 are common. Quantitative calculations of the sputtering yield in this regime are based on the idea of a collision cascade (a primary recoil target atom makes collisions which produce secondaries which in turn produce tertiaries, and so on), and predicted yields typically agree with experiment to within a factor of 2 or 3.

At higher-incident ion energies, in the hundreds of keV per AMU to several MeV per AMU range, for example, the collision cascade predictions for the sputtering yield are small compared with unity. However, recent experiments using high-energy ion beams have shown that for some materials very large sputtering yields ($S \sim 4,000$ in some cases) exist when the collision cascade prediction is $S \ll 1$.

At Bell Laboratories, Walter Brown and his collaborators^{1,2} have shown that H and He in the 100 keV per AMU range can sputter H₂O ice at rates up to three orders of magnitude faster than those predicted according to the low-energy collision cascade calculations. Moreover, at temperatures below ~ 100 K, the yield is independent of temperature, and appears to be proportional to the square of the electronic stopping power, $(dE/dx)_e^2$. This is to be contrasted with the energy dependence of the collision cascade, which is given by the nuclear stopping power $(dE/dx)_n$ describing the energy loss of the projectile to elastic atomic collisions. Evidently energy is transferred from the initial electronic excitation into kinetic

energy of atoms, some of which escape as sputtered particles. Brown and his colleagues favour an explanation based on the ion explosion model. Neighbouring atoms, partially stripped of their electrons by the passage of the projectile, acquire kinetic energy through their mutual Coulomb repulsion. Such a picture leads essentially to a $(dE/dx)_e^2$ prediction for the energy dependence.

An early prediction³, that in certain dielectrics the ion explosion mechanism would lead to an anomalously high sputtering yield essentially proportional to $(dE/dx)_e^2$, led to a series of experiments in the Kellogg Radiation Laboratory at Caltech by Joseph Griffith⁴ and Elizabeth Seiberling⁵. These workers utilized beams of heavy ions (¹⁶O, ¹⁹F, ²⁰Ne) to bombard a dielectric target, UF₄. Sputtered ²³⁵U atoms were collected on mica, fissioned with thermal neutrons, and the resulting fission damage tracks in the mica counted to determine the yield S . This sensitive technique⁶ has allowed the Caltech group to determine the energy spectrum of sputtered particles as well as the dependence on bombarding energy. They discovered that, for ¹⁹F on UF₄, the excitation curve had a much steeper dependence on energy than $(dE/dx)_e^2$, and that the energy spectrum of sputtered particles was essentially thermal at a temperature of 3,500K. To explain their data Seiberling and co-workers have proposed a 'thermalized' ion explosion model, wherein Coulomb energy of neighbouring ion pairs is transformed into random thermal kinetic energy of atomic motion of particles located along the ion path. The model correctly gives the strong energy dependence of the yield, as well as the energy spectrum.

P K Haff and L E Seiberling are in the W K Kellogg Radiation Laboratory, California Institute of Technology, Pasadena

1. Gilbert, W. *Nature* 271, 501 (1978)
2. see Molgaard, H. *Nature* 286, 657 (1980)
3. Lin, C P, Tucker, P W, Mushinski, F J & Blattner, F R. *Science* 209, 1348 (1980)
4. Wall, R. *et al.* *J. supramolec. Struct. & cell Biochem. Suppl.* 5, abstract 5 (1981)
5. Moore, R W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press)
6. Rabbits, T J. *J. supramolec. Struct. & cell Biochem. Suppl.* 5, abstract 30 (1981)
7. Alt, F W. *et al.* *Cell* 20, 293 (1980)
8. Rogers, J. *et al.* *Cell* 20, 303 (1980)
9. Early, P. *et al.* *Cell* 20, 313 (1980)
10. Corte, G. *et al.* *Scand. J. Immunol.* 9, 141 (1977)
11. Maki, R. *et al.* *Cell* (in the press)

The discrepancy between the energy dependence of S , as measured by the two groups, does not necessarily mean that one is right and the other wrong. The assumptions underlying the thermalized ion explosion picture may breakdown for H and He bombardment where the ionization per unit length along the projectile track is considerably smaller than in the UF_4 experiments involving heavy ion projectiles.

Regardless of the precise nature of the mechanisms involved, enough has been learned to know that anomalous sputtering is probably quite important in the Solar System. Extrapolating Pioneer spacecraft data and the Bell Laboratories H_2O sputtering results, Lanzerotti and colleagues⁷ pointed out that over geological time significant erosion can be expected to have occurred on the icy satellites of Jupiter. They argued that up to 1 km of ice may have been lost to space, that interplanetary ice grains would be destroyed in tens of thousands of years by sputtering by the solar wind⁸, and that the ice particles comprising the rings of Saturn may be substantially affected by magnetospheric proton bombardment⁹. Work at Yale and Caltech has focused on the rate of erasure by sputtering of surface features, such as meteorite craters¹⁰, on the sputter-induced formation of planetary atmospheres¹¹, and on the albedo changes of icy satellite surfaces due to differential sputtering of ice and embedded silicate grains^{12,13} (which do not exhibit anomalous sputtering in the energy range considered).

It is clear that there are a large number of potential applications to problems in planetary science. The range of experiments reported so far, however, is relatively small, so, with theoretical ideas requiring further development, much uncertainty is attached to calculations of planetary processes. The Caltech energy spectrum suggests that energies of sputtered particles peak in the 0.3 eV range. This is substantially less than the energy required by, for example, an O atom to escape into space from Ganymede. Will H_2O exhibit a similarly soft spectrum? Probably it will, but until one can measure, or know enough of the mechanisms to estimate reliably, such quantities, it will be difficult to assess with confidence the detailed response of satellite surfaces to ion bombardment. □

Spinning asteroids

from David W. Hughes

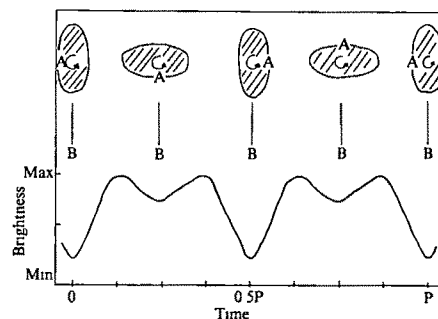
ASTEROIDS spin with periods which usually lie between 5 and 12 hours although they can vary from 4 to 50 hours. The values are obtained from observations of fluctuations in brightness. Large variations are caused by gross changes in the projected cross-sectional area as seen from Earth (see the figure) and smaller variations by changes in the surface properties. Over longer time periods, of the order of a year or more, the light curve shape alters according to variations in the aspect angle — the angle between the line of sight and the asteroid's spin axis. Ideally these long-term variations should lead to an estimation of the asteroid's shape.

Problems arise when interpreting the results because too few asteroids have been observed and observational selection effects must be removed. These problems have been tackled by E. F. Tedesco (Lunar and Planetary Laboratory, University of Arizona) and V. Zappalá (Osservatorio Astronomico di Torino, Italy) in their recent paper in *Icarus* (43, 33, 1980).

About 2,150 asteroids are known and of these 321 have known spin periods. Tedesco and Zappalá divide these into two groups — 187 are members of families and the rest are main belt non-family asteroids. Previous studies in 1973 and 1979 were based on a total sample of only 64 and 182 asteroids respectively, so from this aspect alone Tedesco and Zappalá's work is a considerable improvement.

There are several fundamental selection effects in asteroid light curve data. Two main asteroid classes exist — C types which have a low albedo and are carbonaceous and S types which have a higher albedo and are siliceous or stony-iron in composition. Selection favours asteroids of high apparent brightness, which is equivalent to favouring the largest and nearest and the ones with highest albedo. Unfortunately the C to S ratio varies both as a function of size and distance from the Sun. Apollo and Amor asteroids are also unusual. These have orbits which cross those of Earth and Mars. Their mean rotational brightness amplitude is about 0.5 magnitudes while comparable-sized main belt asteroids have amplitudes usually less than 0.2 magnitudes. Members of the Appolo/Amor group are clearly more irregular and probably have a different evolutionary history to main belt asteroids. Another selection effect is inherent in the photographic photometry used to measure small asteroids. Usually observing periods of only 6 hours can be employed and variations of less than 0.2 magnitudes cannot be detected. Sixty per cent of the observed asteroids do not reveal their spins, indicating that the spin period is too long or the asteroid too regular. A further complication results from the fact that

many small asteroids were specifically selected for observation simply because they were members of Hirayama families. The family members have similar orbits and probably resulted from the fragmentation of a single parent. Removing the selection effects reduces the original data set of 321 to a final set of 134.



A theoretical light curve for a spinning asteroid showing how the brightness varies as a function of time. P is the period. The spin axis is perpendicular to the paper and the asteroid is viewed from B. Position A represents a fixed place on the asteroid.

Tedesco and Zappalá find that there is no clear-cut trend towards faster rotation for smaller objects. This is contrary to the conclusions of previous researchers who regarded collisional fragmentation as an important contributor to the modification of spin rates. While no definite relationship exists, the authors stress that the distribution is not entirely random. Asteroids larger than 175 km have a mean rotation period near 7 hours compared with a 10-hour period for smaller ones.

Considering asteroids with sizes between 75 and 175 km they found that the S asteroids have a mean period of 9.4 hours whereas the C asteroids have a period of 11.6 hours. The difference between these values is only 1.6 times the standard deviation and the authors conclude that this is too small and thus the spin rate is not a function of type.

How do asteroid shapes vary as a function of size and compositional type? It has been shown that the spin axes of asteroids are fixed in space and that it only takes about one million years to align the spin axis with the axis of maximum moment of inertia. Studies also indicate that asteroid surfaces generally lack variegation in colour or albedo. Unfortunately, shapes cannot be deduced directly from light curve amplitudes. Tedesco and Zappalá conclude that the very largest asteroids (diameter $D > 300$ km) have very low-amplitude variations (0.11 ± 0.02 magnitude). There seems to be

David W. Hughes is in the Department of Physics, University of Sheffield.

1. Brown, W. L. *et al.* *Nucl. Instrum. Meth.* 170, 321 (1980).
2. Brown, W. L. *et al.* *Phys. Rev. Lett.* 45, 1632 (1980).
3. Haff, P. K. *Appl. Phys. Lett.* 29, 473 (1976).
4. Griffith, J. E. *et al.* *Rad. Effects* 51, 223 (1980).
5. Seiberling, L. E. *et al.* *Rad. Effects* 52, 201 (1980).
6. Gregg, R. & Tombrello, T. A. *Rad. Effects* 35, 2113 (1976).
7. Lanzerotti, L. J. *et al.* *Geophys. Res. Lett.* 5, 155 (1978).
8. Lanzerotti, L. J. *et al.* *Nature* 272, 431 (1978).
9. Cheng, A. F. & Lanzerotti, L. J. *J. geophys. Res.* 83, 2597 (1978).
10. Haff, P. K. *et al.* in *Proc. 10th Lunar Planet. Sci. Conf.* (Lunar Planet. Inst., Houston, 1979).
11. Watson, C. C. *et al.* in *Lunar Planet. Abstr. XII*, in the press.
12. Haff, P. K. in *Proc. Symp. on Thin Film Interfaces and Interactions* (eds Baglin, J. E. E. & Poate, J. M.) (The Electrochemical Society, Princeton, New Jersey, 1980).
13. Conca, J. in *Lunar Planet. Abstr. XII* (in the press).

an abrupt increase in amplitude as the diameter decreases below 200 km (to about 0.21 ± 0.01). Even so, the shapes of smaller asteroids do not become significantly more elongated with decreasing size.

A search for possible collision-induced effects revealed that asteroids found in the zones of the asteroid belt which have low number densities do seem to have a higher value for the amplitude variation. Unfortunately the relative numbers of the specific asteroid types vary as one moves from zone to zone and this further complicates the picture.

When it comes to asteroid families it seems that the smaller members of the families have higher spin rates than normal. Larger family members ($75 < D < 175$ km) are virtually indistinguishable from non-family asteroids.

The significant conclusion of this investigation is that there is a pronounced

change in asteroid rotational properties at about 200 km. Asteroids larger than this seem to be primordial (at least in the sense that they are not remnants of a larger disrupted parent body), nearly spherical and probably have prograde rotation. Smaller asteroids have random spin axis orientations and the population contains a higher proportion of slowly rotating objects. It is interesting to note that the dividing diameter of 200 km is strongly evident when the frequency-size relationship of asteroids is investigated. The mass distribution index of asteroids changes at this diameter.

Future progress in the study of asteroid spin statistics and the effect of collisional evolution depends on the acquisition of more data. The picture will be a lot clearer if the number of asteroids with accurately measured spin periods could be increased by a factor of two or three. □

Production and processing of prolactin

from Alan D B Malcolm

THE pituitary hormone prolactin (PRL) was for many years something of an enigma. Its similarity to growth hormone both chemically (human PRL and human GH have 40 per cent of their amino acids in common) and biologically (in lower mammals PRL acts as a growth hormone) meant that its very existence in humans was disputed. Advances in protein purification and radioimmunoassay have now resolved this question and we are beginning to understand something about the factors controlling its synthesis and breakdown. Controls seem to exist at many stages including transcription, RNA processing, translation, preprolactin processing and secretion together with nuclease degradation of RNA and proteolytic attack on the various polypeptides.

The mRNA for rat PRL was obtained by extracting the poly(A)-containing RNA from the pituitary after stimulation of the rat with diethylstilboestrol (Gubbins *et al* *Nucleic Acids Res* 6, 915, 1979). This RNA was used to produce cloned copy DNA, which has since been used to identify genomic PRL sequences in a rat genome 'library'. The PRL gene contains four intervening sequences (Chien and Thompson *Proc natn Acad Sci U S A* 77, 4583, 1980), one of which occurs in the 5'-untranslated region. The sequence of one of the genomic clones (Gubbins *et al* *J biol Chem* 255, 8655, 1980) shows that some of the intron sequences are repeated elsewhere in the rat genome. It is interesting that although the coding lengths of PRL and GH are similar and they both have four introns, the lengths of the introns are much

greater in PRL than in GH. The total length of the PRL 'gene' is therefore 10 kilobases compared with only 2.1 for GH.

Maurer and his colleagues (*J biol Chem* 255, 2243; 1980) have approached the problem of transcription and processing of PRL RNA by isolating RNA from the nucleus and cytoplasm of pituitary cells. The RNA was then electrophoresed through an agarose gel and transferred to diazobenzylxymethyl paper. Hybridization of the now immobilized RNA to ^{32} P-labelled cDNA revealed the presence of five PRL bands (7.0, 6.4, 3.8, 1.7 and 1.0 kilobases) in nuclear RNA but only one (1.0 kilobase) in cytoplasmic RNA. Presumably the different nuclear RNA species arise from ordered removal of the introns.

The secretion of all the pituitary hormones is controlled by the hypothalamus. In the case of PRL, dopamine is probably an important inhibitory factor and many dopamine agonists, such as the ergot alkaloids ergocryptine (EC) and its synthetic congener bromoergocryptine (BEC), will reduce overall PRL production. EC is used clinically to reduce PRL levels in spite of slightly unpredictable side effects (*New Engl J Med*, 302, 749, 1980).

To investigate the stage at which EC regulates PRL production, Maurer (*J biol Chem* 255, 8092, 1980) has studied the decrease in both *de novo* PRL synthesis

and PRL mRNA in cultured pituitary cells after treatment with EC. The very good correlation between the two suggests that EC acts either at transcription, RNA processing or RNA breakdown. It is slightly odd that the decreases occur over 20–100 hours which suggests that the mRNA may have a fairly long half life. Interestingly though, EC affects PRL specifically (no other pituitary hormones are affected) and is 100 times more effective than is dopamine.

This picture is, however, somewhat confused by the finding that BEC increases the rate of breakdown of the hormone itself (Dammes and Rudnick *J biol Chem* 255, 2776, 1980). Maurer (*Biochemistry* 19, 3573, 1980) suggests that BEC blocks PRL secretion by the cells and the subsequent accumulation inside the cells results in increased proteolysis by lysosomal enzymes. The fact that chloroquine (which reduces the rate of appearance of lysosomal enzymes in the cytoplasm) partially reverses this effect of BEC favours such an interpretation.

Hormone production can also be prevented by incubating the pituitary cells with hydroxynorvaline — an analogue of threonine which can be incorporated into protein. Normal processing of preprolactin to PRL involves cleavage at the bond between residues 30 (normally threonine) and 31. Apparently the processing enzyme cannot hydrolyse this bond when threonine is substituted by hydroxynorvaline because the prehormone accumulates in the analogue-treated cells (Hortin and Boime *J biol Chem* 255, 7051, 1980).

Rossier and colleagues (*Proc natn Acad Sci U S A* 77, 666, 1980) have shown that stress induces PRL synthesis and they believe that this arises from the stress-induced production of opiate peptides which are known to increase PRL levels.

More recent results (Denef *et al* *Nature* 285, 243; 1980) have suggested that the action of dopamine is not quite as simple as was previously thought. Although high levels ($\geq 1 \mu\text{M}$) of dopamine do indeed reduce PRL, at much lower levels ($\leq 1 \text{ nM}$) dopamine actually stimulates PRL production.

The nature of the interactions between prolactin and the steroid hormones is still not clear. Although oestradiol seems to stimulate PRL production, its metabolite, 2-hydroxyoestrone, produces a rapid decrease in PRL levels (Fishman and Tulchinsky *Science* 210, 73, 1980). The hydroxylating enzyme is inhibited by the opiate agonists which stimulate PRL release and is activated by antagonists which suppress PRL. This means that *in vivo* the opiates may affect PRL levels by their effect on the enzymes which metabolize the steroid hormones.

Clearly we still have some way to go in our attempts to understand how the levels of this interesting molecule really are regulated. □

Alan D B Malcolm is in the Biochemistry Department at St Mary's Hospital Medical School, London.

Tropical soils and agrotechnology transfer

from Ian Smalley

MORE food needs to be produced in the third world, in the countries which are grouped in the tropical regions. But the soils of the tropics often have a 'variable charge' character which makes them sufficiently different from most soils of the temperate regions that research in the developed world can not be directly applied. The transfer of agriculture technology between the developed and the developing world can thus encounter fundamental problems.

Various aspects of variable charge soils, and in particular the problems of classification and of transfer of agricultural technology, were discussed at a recent meeting* at Massey University in New Zealand. In variable charge soils the electrical charge on the particle surfaces, both in magnitude and sign, can vary greatly with changes in the pH, ionic strength and composition of the soil solutions. The variable charge nature of the soils arises from the abundance, in the clay fraction, of minerals having amphoteric (positive or negative) surfaces. Prominent among such minerals are the oxides and hydrous oxides of aluminium and iron which may be crystalline, paracrystalline or poorly ordered. The 1:1 type layer silicates, such as halloysite and kaolinite, also show amphoteric properties, associated with their relatively large exposed crystal edge surface. On the other hand, the negative surface charge of 2:1 clay minerals arises from extensive isomorphous substitution within the structure and hence is largely independent of pH. Thus soils rich in micaceous and similar minerals are said to be of constant or permanent charge.

Besides having amphoteric surfaces, clay-size mineral constituents with variable charge generally have a low capacity to adsorb or retain nutrient cations in the range of pH and ionic conditions which prevail in natural soil systems. For this reason, these constituents have been grouped together and referred to as 'low activity clays'. Thus, soils rich in low activity clays are, in general, naturally infertile and this causes a major problem in tropical agriculture. Although the amphoteric nature of some soils and soil constituents has been recognised for a long time, it is only relatively recently that its importance, in terms of soil processes and management, has come to be appreciated. This is because soil research, as carried out in the developed countries of Europe and North America, has largely been directed to understanding the behaviour of temperate soils, many of which are of the

constant charge type. The variable charge constituents, if they were not altogether ignored, were often regarded as contaminants and removed from the system.

A common constituent of volcanic variable-charge soils is the clay mineral allophane (now widely believed to have a spherical structure see *News & Views, Nature*, 281, 339, 1979). Recent studies on allophane in New Zealand soils were reported by R. Parfitt (NZ Soil Bureau). Allophane has been defined as a series of naturally occurring hydrous aluminosilicate clays, characterised by short range order and by the predominance of Si-O-Al bonds, and it appears to be made up of hollow spherules 3.5–5.0 nm in diameter. Spherules with a molar ratio close to Al/Si = 2.0 from Mt Egmont andesitic pumice have the proto-imogolite infrared spectrum. This, together with X-ray fluorescence, pyridine adsorption and silicate analysis data, indicate that the wall of the allophane spherule is made up of imogolite segments. The fine tubular crystals of imogolite are also found in New Zealand soils.

Allophane (Al/Si = 2.0) occurs in several different environments in New Zealand. It is the dominant clay mineral in an Egmont soil formed from andesitic ash, in a Taupo soil formed from rhyolitic Taupo pumice, and in a Maoroa soil formed from rhyolitic ash. In rhyolitic ash, allophane (Al/Si = 2.0) is the dominant clay mineral where the annual rainfall exceeds about 1,600 mm. Where the annual rainfall is less than 1,600 mm, allophane and halloysite are the dominant clay minerals, with allophane decreasing and halloysite increasing in older tephra. Halloysite is the dominant clay mineral in a Kereone soil, where the rainfall is 1,200 mm.

Allophane soils present a problem for mechanical analysis and clay recovery because at field pH or in acid media the positively charged allophane flocculates with any crystalline clay present by charge interaction, and in alkali or calgon solution it is self-flocculating by some unknown mechanism. Studies of Scottish allophanic subsoils by H. J. Fullerton (University of Glasgow) showed that these behave like New Zealand allophane soils in that they flocculate by charge interaction when vibrated in water and continue to flocculate at all alkali pH values. Unlike allophane soils, however, they deflocculate in calgon solution. Alkaline flocculation appears to be not a rare property confined to allophane, but a common property of disordered clay surfaces.

The major technology transfer project discussed, the Benchmark Soils Project, is being applied to three soil types across the tropical world. Three different soil families have been chosen for experimentation: the

clayey, kaolinitic, isohyperthermic Tropeptic Eutrustox (the terminology of the USDA-developed 'soil taxonomy' system is used), the thixotropic, isothermic Hydric Dystrandepts, and clayey, kaolinitic, isohyperthermic Typic Paleudults. These soil families represent three different agroclimatic zones: dry and warm (the Eutrustox), moist and cool (the Dystrandepts), and moist and warm (the Paleudults). The soils all have distinct variable charges which influence soil management and utilization. G. Uehara (University of Hawaii) reported that technology transfer in 15 separate areas was being attempted, among these were selection of tree crops for windbreak, fuelwood and erosion control, susceptibility of salinization, and soil workability and trafficability.

The Benchmark Soils Project comprises two companion research contracts awarded to the Universities of Hawaii and Puerto Rico in 1974 and 1975 by the US Agency for International Development. The primary objective of the project is to test the hypothesis that agro-production technology, particularly soil management experience, can be transferred among tropical countries on the basis of soil families as defined in the system of soil taxonomy. The project has established a research network that includes 23 experimental sites in Brazil, Cameroon, Hawaii, Indonesia, the Philippines and Puerto Rico. The basic research strategy is to select soils belonging to the same soil families in widely separated geographical areas, subject these soils to identical and standardized experiments, and correlate the plant yields. The transfer experiments are complemented by pragmatic management experiments and variety trials intended to expand the basic knowledge of crop production for each soil family.

R. L. Fox (University of Hawaii) pointed out that in their attempts to replace the humus theory of plant nutrition with the inorganic nutrient theory, early naturalists assumed that it was only necessary to replace the nutrients removed in the harvest. Phosphate reversion, K fixation, denitrification, soil erosion and leaching of nutrients were not sufficiently recognized. We now know that a soil fertility programme that is content with returning nutrients exported in the harvest is not sufficient at all. This fundamental mistake has been repeated almost endlessly during the past century — most recently in the tropics. Nowhere are the consequences likely to be so acute as on the weathered, variable charge soils which are abundant there. □

*'Soils with Variable Charge' Conference. Commissions 4-7 of the International Society of Soil Science, 11-18 February, 1981. Conference Book (ed. B. K. G. Theng) publ. NZ Soc. Soil Sci. 1981.

Ian Smalley is with the New Zealand Soil Bureau, Lower Hutt, New Zealand.

Interfering plants

from Peter D. Moore

IN his classification of competitive strategies, Miller (*Adv ecol Res* 4, 1, 1947; and *Brookhaven Symp Biol* 22, 63, 1969) recognizes two major categories, exploiters and interferers. The interferer has a large body size, a long generation time and is relatively independent of physical environmental factors. Rather than evolving efficient mechanisms for resource exploitation, they substitute space for the resource and succeed by means of their capacity for the physical exclusion of potential competitors from the resource. The concept was developed largely with certain vertebrate animals in mind, especially those exhibiting territorial behaviour.

There are also some members of the plant kingdom which deserve the title 'interferers'. In particular, those which assume dominance in stable ecosystems have characteristics which correspond to those of the interferer. The substitution of space for a resource may be achieved purely by physical growth, or may be achieved by more subtle means. The production of chemical inhibitors which curtail the growth of plant individuals of either the same or another species is one such mechanism. Even more subtle is the process described by Rice and Pancholy (*Am J Bot* 59, 1033, 1972, see also *Nature* 254, 184, 1975) in which toxins from litter of grassland climax dominants is considered to inhibit the process of nitrification (the conversion of ammonium ions to nitrate in the soil) by suppressing the activity of the soil microbes *Nitrobacter* and *Nitrosomonas*. Such a strategy would reduce the rate of nitrogen recycling and thereby slow down further ecosystem growth and vegetation succession.

This attractive and sophisticated

interference technique has, however, received little supporting evidence over the past few years. It is difficult to distinguish between direct chemical suppression and rather more complex interactions involving the depletion of other necessary resources (for example, phosphorus) or the overall reduction of pH, both of which could influence the rate of nitrification. Jordan, Todd and Escalante have tried a technique of manipulative experimentation in the Amazonian rain forest (*Oecologia* 39, 123, 1979). They added calcium carbonate to the thick root mat on the forest floor and checked the leachate for nitrogen losses. They observed an enhanced loss of nitrogen, but it was mainly in the form of ammonium ions, so the treatment had not stimulated nitrification. They also noted that the pH of the drainage water was still as low as 4.5 and thus concluded that it was the low pH of the soil which was inhibiting the activity of the nitrifying bacteria. However, they conceded that, since the leachate had the colour and consistency of tea, it could be the tannin from the leachate which was responsible for the suppression.

The deciduous forests of Missouri have provided opportunities for further analysis of the curtailment of nitrification. Here Lodhi (*Am J Bot* 64, 260, 1977, 65, 340, 1135, 1978) has shown a heterogeneous pattern in the nitrogen levels of soils under different canopy tree species. Ammonium ion concentrations were invariably higher than nitrate levels, the degree of excess being related to canopy species. Also, there was an inverse relationship between am-

monium ion concentration and the density of nitrifying bacteria in the soils. Under some trees, such as American elm (*Ulmus americana*) and sycamore (*Platanus occidentalis*), the soils were of neutral pH, so one cannot resort to a low pH explanation for the suppression of nitrification at all these sites.

Lodhi has now backed up his contention that polyphenolic compounds can be involved in these woodland inhibition observations by his studies with Killingbeck (*Am J Bot* 67, 1423, 1980) on the ponderosa pine stands of North Dakota. Here the soil pH is neutral to slightly alkaline (mean pH 7.25 in the A horizon and 7.75 in the C horizon). Once again, the ammonium ions considerably exceed nitrate ions and the population of *Nitrosomonas* and *Nitrobacter* were low. He then added various extracts of ponderosa pine needles and bark to experimental cultures of these microbes and found that maximum inhibition of their activity (93 per cent) was achieved with aqueous extracts of the needles or with mildly alkaline soil hydrolysates. Both of the extracts were rich in condensed tannins, which probably accounted for the inhibition.

Nothing is known of the mechanism of nitrogen uptake by ponderosa pine but Lodhi speculates that there would be considerable ecological advantages in its taking up ammonium rather than nitrate ions. In this way leaching losses of nitrogen would be reduced, energy would be conserved and the tree would benefit in terms of competition for the soil nitrogen resources. More data is required before such speculations can be confirmed but, should they prove correct, *Pinus ponderosa* will effectively fit the Miller criteria defining an interferer. Whilst satisfying its own requirements, it denies other organisms access to the soil nitrogen resource by allelochemic inhibition of nitrification. □

Peter D. Moore is in the Department of Plant Sciences, University of London, King's College



100 years ago

THE NEW MUSEUM OF NATURAL HISTORY

The great terra-cotta building facing Cromwell Road, South Kensington, and occupying the site of the old 1862 Exhibition, about which for the past twelve months public curiosity has been raised, is about to draw up its blinds, and offer a part of its extensive galleries for inspection on Easter Monday.

It is no secret that for thirty years past the accommodation in the British Museum, Great Russell Street, Bloomsbury, "both for man and beast," had become too restricted, and the necessity for a larger building was keenly felt. As usually happens in such cases, the most adventurous and energetic officer was the first to obtain for his department what he

required, namely, *more room*.

Sir A. Panizzi (then Keeper of Printed Books) projected, shortly after the 1851 Exhibition, his scheme for a great central Reading-room and Library, and some years later on, the Department of Antiquities, represented by Mr C. T. Newton, C.B., also obtained an addition to its galleries on the western side, and still more recently on the southern side, next to the great Entrance Hall.

Great praise is due to Mr Bond, the present Principal Librarian, for putting an end to the use made of the fine colonnade in front of the British Museum, which for twenty-five years was blocked by antiquities covered in with a row of extremely unsightly and incongruous wood and glass sheds. These are now happily removed. The Department of Prints and Drawings lacking a gallery, obtained possession of the "King's Library" floor on the east side for an exhibition space, and even the conservative Coin Department likewise laid out for the public a few show-cases here of coins and medals.

But, like the clothes of the rising son, in the old caricature, the collections everywhere had outgrown their receptacle, and none more so than the departments of Natural History. Scientific men were however not unanimous, and a fierce controversy was carried on in 1858-59 as to the relative merits of enlargement on the old site, or dismemberment. Finally, after a Committee of the House of Commons

Erratum

In *A tropical Volute shell and the Icarus syndrome* from Jonathan C. Howard (*News and Views* April 9, 290, 441, 1981) 'simulate' was incorrectly printed as 'stimulate' on p442, column 3, line 37. The correct sentence reads "C. H. Waddington⁸ pointed out as long ago as 1942 that genetic modifications to adaptable systems can simulate the inheritance of acquired characters."

had taken evidence upon the subject, the removal of the Natural History Collections was decided upon by the Government.

But the death of the Prince Consort, the delay of the House of Commons to vote the necessary funds, the retirement of Sir A. Panizzi from the post of Principal Librarian, the discussion of rival plans, the inevitable delays about the completion of any Government building, caused twenty years to pass before the plans of the chosen architect, Mr Alfred Waterhouse, were realised in a solid and material form.

From *Nature* 23, 14 April, 549, 1881

REVIEW ARTICLE

Dynamics, diffusion and geomorphological significance of tidal residual eddies

J. T. F. Zimmerman

Netherlands Institute for Sea Research, PO Box 59, Texel, The Netherlands

Many shallow tidal areas show the occurrence of ebb and flood surpluses, organized in cell-like structures or 'tidal residual eddies'. They seem to arise from a nonlinear transfer of vorticity from the oscillating tidal field to the mean field, the irregular structure of the bottom topography or the coastline acting as a catalyst to produce the necessary vorticity gradients. The dynamical characteristics of the eddies assist our understanding of the geomorphology of the sea bed.

VAN VEEN'S review¹⁻³ of the hydrography of the Straits of Dover and its influence on the morphology of the Dutch coast stimulated much research on the connection between tidal currents and a variety of coastal geomorphological features. Much attention has since been given to the occurrence of linear or parabolic sand ridges in tidal areas, such as the Flemish Banks in the Southern Bight of the North Sea. We now know that nearly all shallow tidal areas where oscillating currents exceed some critical speed (0.5 m s^{-1}) and where sand is present, are sites at which these ridges can be found (see Fig. 1)⁴. Tidal currents in the vicinity of these ridges are generally asymmetrical in that at some places more water seems to be displaced in the direction of the flood current (a flood surplus) whereas the reverse occurs elsewhere (an ebb surplus). These surpluses, called residual currents, are often organized in cell-like structures or 'residual eddies' (see Fig. 2). Yet, although these features have long been recognized, the underlying physical dynamics of residual eddies in connection with tidal sand ridges have only recently become understood.

The dynamics of residual eddies occurring in geomorphological conditions other than tidal sand ridges, such as the generation of residual eddies by headland promontories or by

narrow tidal inlets along the coast and the organization of ebb and flood surpluses by the gross geometry of half-enclosed tidal embayments, have also been extensively investigated. Although, at first sight, it seems that different mechanisms are at work in each geomorphological setting, it is shown here that the formation of residual tidal circulation can in each case be understood as a transfer of vorticity from the tidal oscillating current velocity field to the mean (residual) field, though the way in which tidal vorticity and tidal vorticity gradients are produced in the first instance differ.

Although we now have a good understanding of the fluid dynamical response of an irregular sea bed or coastline to the tidal current regime, we are still far from having a satisfactory theory for the interaction of tidal currents and morphological structures concerned, that is, a complete explanation of how the morphology is both being shaped by, and gives shape to the current velocity field. However, the recently developed physical description of tidal residual circulation gained by observations in the field, by hydraulic or numerical modelling and, principally, by theoretical analysis may provide a basis for such a theory and may also help our understanding of the contribution of residual eddies to diffusive transport processes in tidal areas.

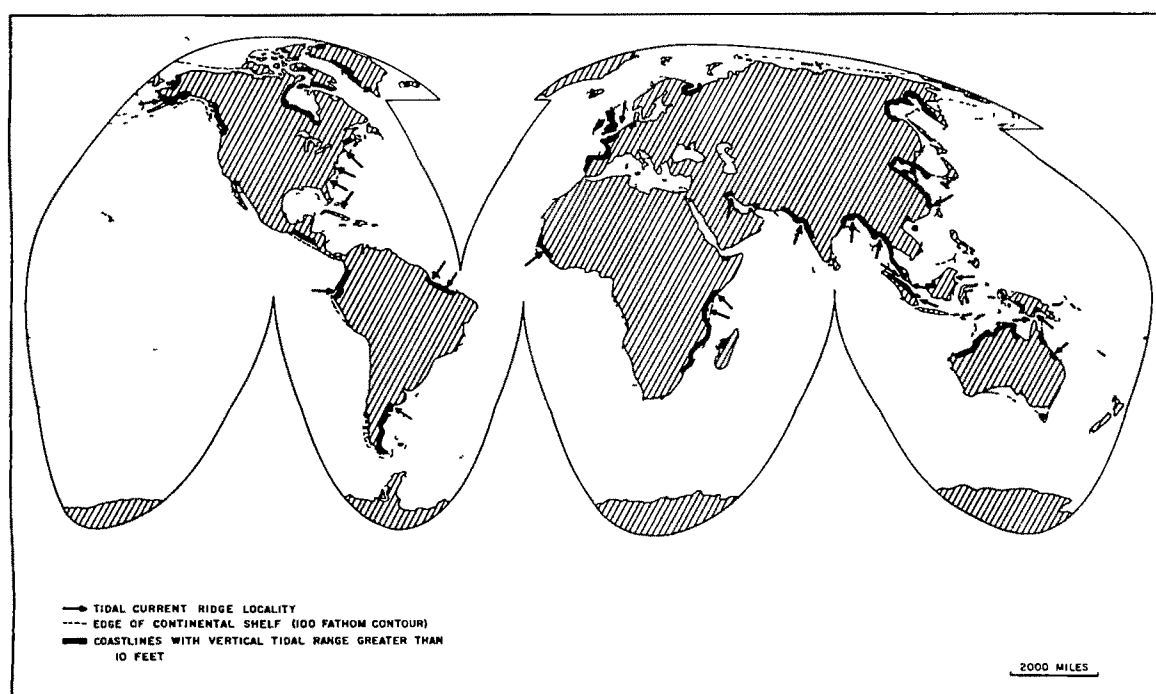


Fig 1 The worldwide distribution of tidal current ridges according to ref. 4

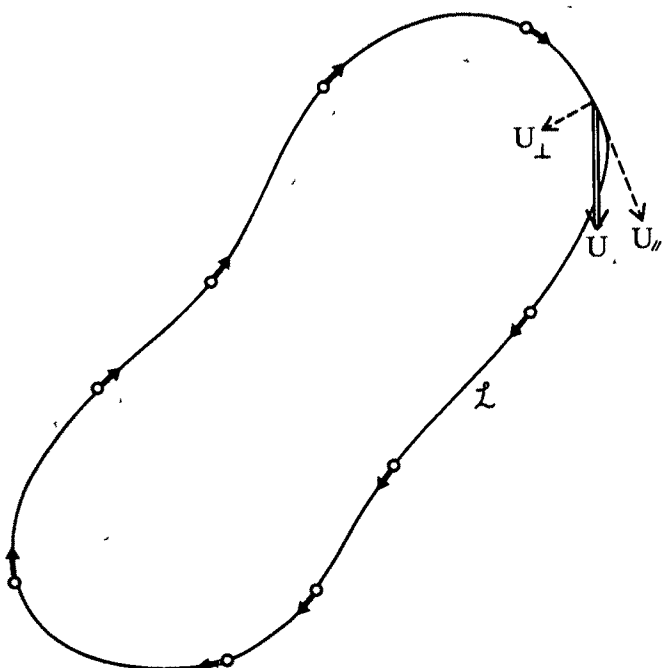


Fig. 2 A string of current meters produces, after averaging the instantaneous velocity $\mathbf{u}(t)$ over a tidal cycle, residual current vectors ($\rightarrow$) tangential to which a residual streamline $\mathcal{L}$ is drawn. The instantaneous velocity vector can be decomposed into a component $\mathbf{u}_{\parallel}$ tangential to $\mathcal{L}$ and a component $\mathbf{u}_{\perp}$ perpendicular to $\mathcal{L}$.

Vorticity dynamics

The relevance of vorticity to residual eddy dynamics can readily be appreciated if the concept of circulation, which can easily be visualized and has a close connection with vorticity, is introduced as a measure of eddy strength.

Figure 2 shows a hypothetical string of current velocity meters in a quasi two-dimensional (horizontal) tidal velocity field. A simple time-dependent velocity field is assumed. At each position the velocity has two components: one which is strictly periodic in one of the basic tidal frequencies, σ , and the other is a time-independent residual. The current meters are supposed to produce non-vanishing residual-velocity vectors (tidal averages), tangent to which a curve $\mathcal{L}$ is drawn. The curve itself represents a residual streamline, its shape suggests the presence of a residual eddy in the field. The instantaneous velocity vector at each point of $\mathcal{L}$ is

$$\mathbf{u}(t) = \mathbf{u}_{\parallel}(t) + \mathbf{u}_{\perp}(t) \quad (1)$$

where $\mathbf{u}_{\parallel}$ and $\mathbf{u}_{\perp}$ are components tangential and perpendicular to $\mathcal{L}$ respectively. Denoting tidal averaging by $\sim$, we have by definition of $\mathcal{L}$ and in view of the simple time dependence of the velocity field

$$\begin{aligned} \tilde{\mathbf{u}}_{\perp} &= 0, \quad \tilde{\mathbf{u}}_{\parallel} = \tilde{\mathbf{u}} \\ \mathbf{u}_{\perp} &= \mathbf{u}'_{\perp} \sin(\sigma t + \phi_{\perp}), \quad \mathbf{u}_{\parallel} = \tilde{\mathbf{u}}_{\parallel} + \mathbf{u}'_{\parallel} \sin(\sigma t + \phi_{\parallel}) \end{aligned} \quad (2)$$

where $\phi_{\perp}$ and $\phi_{\parallel}$ are arbitrary phase angles. To get some measure of the eddy strength, the circulation, C , around $\mathcal{L}$ is now defined by the contour integral

$$C(t) = \oint \mathbf{u}(t) \cdot d\mathbf{l} = \oint u_{\parallel}(t) dl \quad (3)$$

By Stokes' theorem this integral equals an integral over the surface A enclosed by $\mathcal{L}$

$$C(t) = \iint_A \omega \, dA \quad (4)$$

where ω , the vorticity of the current velocity field, can be interpreted as the 'spin' or twice the angular velocity around a

vertical axis of an infinitesimal fluid element inside $\mathcal{L}$. Now, taking tidal averages, the residual circulation reads

$$\tilde{C} = \oint \tilde{u}_{\parallel} dl = \iint_A \tilde{\omega} \, dA \quad (5)$$

Thus, if residual circulation around a closed curve in a tidal area is positive (anticlockwise or cyclonic) or negative (clockwise or anticyclonic) then, on average, a distribution of negative or positive residual vorticity ('spin') of the fluid elements inside the area exists over the area enclosed by the curve.

To find when tidal regions exhibit areas of non-vanishing tidal average vorticity we need to look more closely at the equation of motion of a quasi two-dimensional homogeneous fluid flow in a rotating frame of reference

$$\frac{\partial \mathbf{u}}{\partial t} = -\{f + \omega\}(\mathbf{j} \times \mathbf{u}) + \nabla(g\zeta + u^2/2)$$

local acceleration =

Coriolis + vortex force + dynamic pressure gradient (6)

$$-\frac{r}{D} \mathbf{u} + \nu \nabla^2 \mathbf{u}$$

(linearized) bottom friction + horizontal turbulent momentum exchange

The local acceleration at a particular position in the field equals the sum of the forces in the right-hand side of equation (6), f denoting the Coriolis parameter (the planetary vorticity), g , the acceleration of gravity, ζ , the departure of the sea surface from its mean position, r , a bottom friction parameter, D , the water depth, and ν , a turbulent viscosity coefficient. Neglecting the last term in equation (6) the equation of motion is readily transformed into an equation for the circulation around $\mathcal{L}$, which reads

$$\frac{\partial C}{\partial t} = \oint (f + \omega) u_{\perp} dl - \frac{r}{D} C \quad (7)$$

The residual circulation is obtained by averaging equation (7) over a tidal cycle

$$\begin{aligned} \tilde{C} &= \frac{\oint \overline{\omega(t) u_{\perp}(t)} dl}{r/D} = \\ &= \text{tidal mean flux of vorticity through } \mathcal{L} \\ &\quad \times \text{relaxation time scale of bottom friction} \end{aligned} \quad (8)$$

Thus a non-zero (residual) circulation around $\mathcal{L}$ is produced if, when averaged over the tidal cycle, there is a net flux of vorticity into the region enclosed by the curve (a net outflux is equivalent to a net influx of vorticity of opposite sign). Such 'pumping' leads to the build-up of positive or negative residual vorticity, ultimately balanced by the dissipation due to bottom friction, or, had we included that effect in equation (7), to horizontal turbulent vorticity diffusion. Note that $\oint \omega u_{\perp} dl$, having dimensions $[L^2 T^{-2}]$, can be interpreted as a stress (momentum flux), often called the 'tidal stress'^{5,6}. Hence the residual vorticity flux can be thought of as the analogue of a time-independent tangential stress exerted along the boundary of the water column inside $\mathcal{L}$, which therefore tends to spin-up.

For a non-vanishing residual circulation we need to both produce vorticity somewhere in the field and achieve a net flux of vorticity through a closed curve over a tidal cycle. Thus, whereas the condition necessary for residual circulation, expressed by equation (8), is now revealed physically, how does such a net flux arise? Mathematically, if $u_{\perp} \propto \sin(\sigma t + \phi_{\perp})$ and $\omega(t) \propto \sin(\sigma t + \psi)$, it is the phase difference $\phi_{\perp} - \psi$ that determines sign and magnitude of the pumping term $\oint \omega u_{\perp} dl$. If $\phi_{\perp} - \psi$ is near to 0 or π that is if $u_{\perp}$ and ω change sign approximately at the same moment in the tidal cycle, then conditions are favourable for the

Table 1 Observed orders of magnitude of residual vorticity

Area	Interpretation measurements supported by	Eddy characteristics			Tidal velocity amplitude (m s ⁻¹)	Eddy type	Refs
		Velocity (m s ⁻¹)	Length scale (10 ³ m)	Vorticity (s ⁻¹)			
a Current measurements							
Kasado Bay Japan	Numerical ⁸ /hydraulic ^{7, 19} model	0.1	2	5 × 10 ⁻⁵	4 × 10 ⁻¹	a	44
Minas Basin Bay of Fundy	Numerical ¹⁰ model	0.2	8	3 × 10 ⁻⁵	1.5	b	45
Portland Peninsula English Channel	Numerical model	0.3	30	10 ⁻⁵	?	b	11, 21
Norfolk sand banks North Sea	Analytical model/bottom topography	0.1	30	3 × 10 ⁻⁶	1?	c	13
Wadden Sea	Bottom topography	0.1	5	2 × 10 ⁻⁵	0.9	a?/c	24, 33
Cook Inlet Alaska		0.2	10	2 × 10 ⁻⁵	1.5	c?	17
Southern Bight North Sea	Analytical model/bottom topography	0.01	40	10 ⁻⁷ –10 ⁻⁶	0.7	c	18
Prototype							
b Numerical models							
Kasado Bay		0.01	5 × 10 ⁻³	2 × 10 ⁻³ (10 ^{-5*})	5 × 10 ⁻³	a	8
Start Point Promontory		0.3	4	<10 ⁻⁴	1	b	12
English Channel		0.2	20	<10 ⁻⁵	1	b	22
Southern Bight North Sea		0.05	15	3 × 10 ⁻⁶	1	c?	5
Southern Bight North Sea		0.1	20	5 × 10 ⁻⁶	0.7	c?	23
c Hydraulic models							
Kasado Bay		0.01	5 × 10 ⁻³	2 × 10 ⁻³ (10 ^{-5*})	5 × 10 ⁻³	a	7, 19
Seto Inland Sea		0.15*	20*	10 ^{-5*}	1	b	6, 20

Values are overall estimates, subject to considerable uncertainty

* Model result rescaled to the prototype, eddy types are in accordance with Fig. 3

buildup of residual vorticity. Figure 3 indicates when such favourable conditions are met, summarizing the different situations dealt with recently. We classify these situations according to the geomorphological background.

(1) Residual basin circulation (Fig. 3a, refs 7–9) is the simplest example of tidal residual circulation and is given by the distribution of ebb and flood surpluses, generated by frictional boundary layers in a semi-enclosed basin. In such a basin the tidal-current velocity diminishes both towards the back wall and, by friction, towards the side walls. By the latter process vorticity is generated, the sign of which is opposite at both sides of the basin and changes with the current direction. As the tidal current decreases towards the back wall, at the left-hand side of the basin, during the flood, more anticlockwise vorticity is brought into the heavily drawn curves near the entrance than is transported outwards near the back wall. During the ebb more clockwise vorticity is brought out, which is equivalent to an influx of anticlockwise vorticity. Thus, averaged over the tidal cycle there is a net influx of anticlockwise vorticity resulting in an anticlockwise residual circulation along the heavily drawn curve in the left-hand side of the basin. The opposite occurs on the right-hand side where a clockwise residual circulation is set up. Thus a flood surplus exists in the middle of the basin and an ebb surplus along the sides.

(2) Headland eddies (Fig. 3b, refs 10–12)—along a coastal promontory the tidal flow is accelerated and decelerated in the current direction and has a maximum near the headland. Concurrently a frictional boundary layer develops diminishing the current velocity towards the coast. Thus vorticity is produced along the coast with its sign alternating between the ebb and flood stages of the tide. As the vorticity is largest near the headland there is a net flux of clockwise vorticity into the upper heavily drawn curve and out of the lower one when the current direction is upwards, during the flood, for example. During the ebb, the situation does not reverse because of a net flux of anticlockwise vorticity out of the upper curve and into the lower curve, which is equivalent to the situation during the flood. Hence there is a net flux of clockwise vorticity into the upper curve and of anticlockwise vorticity into the lower curve, giving rise to a vortex pair of opposite sign at both sides of the promontory.

Note that by placing at the left-hand side of Fig. 3b its mirror image, a quadruple vortex system is obtained which represents

the situation close to a sudden narrowing of a tidal channel, as often occurs near inlets connecting tidal channels and the open sea.

(3) Sand ridge eddies (Fig. 3c/d, refs 13–16)—when a tidal current crosses a submarine sand ridge, the flow is accelerated going up the slope and decelerated going down the slope. Consequently a water column experiences a slightly higher Coriolis force on its shallower side than on its deeper side. Thus a 'Coriolis torque' exists on the column having opposite sign at both sides of the ridge, and, at the same side, at both stages of the tidal period. A torque is also produced by bottom friction, the latter being stronger at the shallower side of a fluid column. Both torques reinforce each other if the ridge is inclined anticlockwise with respect to the tidal current (Fig. 3c), whereas they counteract in the opposite case (Fig. 3d). Taking a curve around the ridge, it is obvious that in case (c) there is a net influx of clockwise vorticity or, equivalently, an outflux of anticlockwise vorticity produced by both torques during the flood (current velocity upwards), assuming that vorticity production is quasi-static, that is the instantaneous sign of vorticity is the same as that of the applied torque. The same occurs during the ebb (current velocity downwards), so that a residual influx of clockwise vorticity produces a clockwise residual circulation around the ridge. In case (d) the same applies to residual circulation related to vorticity production by Coriolis torques but the reverse applies to the residual circulation produced by bottom frictional torques, the latter now giving rise to anticlockwise residual circulation. As both mechanisms counteract, a weaker net circulation may be expected compared with case (c). Of course, this applies to situations in the Northern Hemisphere. In the Southern Hemisphere ridges inclined clockwise with respect to the tidal current would be expected to show reinforcing of both mechanisms.

Scale effects

Are residual eddies of all sizes equally probable or are eddies of a specific horizontal extent preferred dynamically? This question, relevant to the geomorphology of the sea bed, has only previously been answered^{14, 15} for the production mechanisms shown in Fig. 3c, d, that is residual eddies produced by depth gradients in the open sea. Consider a test topography (Fig. 4a) consisting of a regular series of humps and bumps with a height or depth h below the mean depth H and a horizontal wavelength

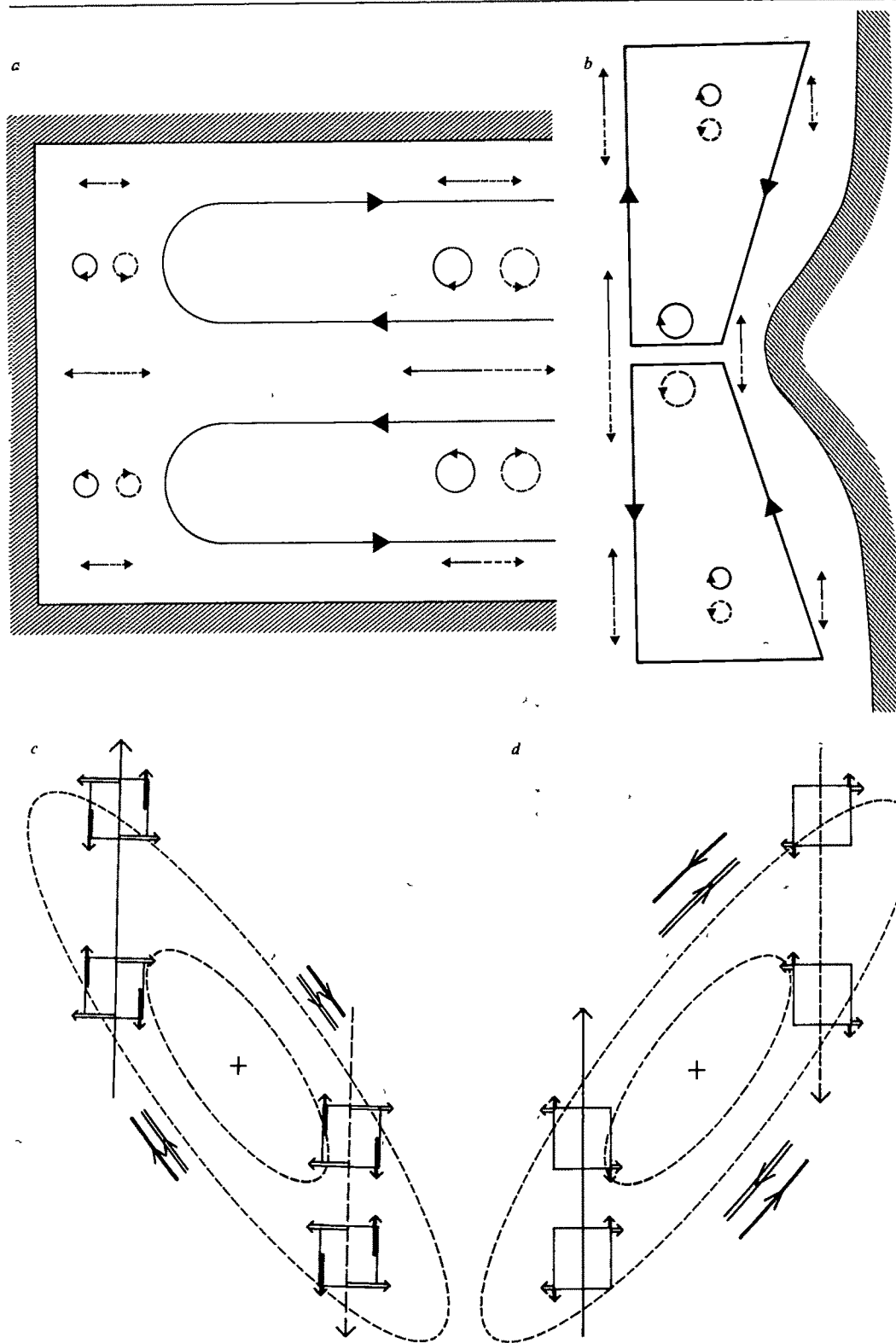


Fig. 3 The basic production mechanisms of residual circulation in tidal currents through sidewall friction (a, b) or through Coriolis and bottom frictional torques (c, d). a, Rectilinear tidal currents in a semi-enclosed basin are shown by solid arrows for the flood period and dashed arrows for the ebb period. The currents decrease both towards the end of the basin and towards the side walls. Vorticity generated by side-wall friction is shown by solid circles for the flood period and dashed circles for the ebb period. The vorticity has opposite signs at both sides of the basin and diminishes towards the back wall. The resulting residual circulation is shown by heavy streamlines. b, Tidal currents, tidal vorticity and residual circulation in the neighbourhood of a headland promontory have the same symbols as in a. Tidal currents are largest near the headland and decrease towards the coast. Tidal vorticity therefore has a maximum near the headland. c, d, A submarine sand ridge is shown by ellipsoidal dashed depth contours, + represents the top of the ridge. Rectilinear tidal currents, solid arrows for the flood period, dashed arrows for the ebb period, cross the ridge, which in c is inclined anticlockwise with respect to the current and clockwise in d. By the presence of bottom depth gradients both the Coriolis force and bottom friction produce torques at a fluid column as shown by the open arrows for the former and solid arrows for the latter. The situation during the flood period is only shown for the left-hand side of the ridge and during the ebb period only for the right-hand side. The bottom frictional torques change sign if the inclination of the ridge changes from anticlockwise to clockwise, whereas the Coriolis torques remain unchanged. The resulting residual circulation around the ridge is shown by solid and open arrows, parallel to the depth contours, the arrows denoting respectively circulation produced by the bottom friction mechanism and by the Coriolis mechanism.

L The tidal current is characterized by a velocity amplitude U and the frequency σ , which produce a length scale, the amplitude of horizontal tidal excursion, l

$$l = U/\sigma \quad (9)$$

Together, tide and topography are now characterized by two non-dimensional parameters, a relative depth h/H and a wavenumber, λ

$$\lambda = \frac{2\pi l}{L} = \frac{\pi \times \text{tidal excursion}}{\text{topographic wavelength}} \quad (10)$$

The ratio v of the residual speed around a curve enclosing a

hump and the amplitude of tidal velocity is^{14 15}

$$v \equiv \frac{\bar{u}}{U} = \frac{h}{H} g(\lambda) \quad (11)$$

The weighting function $g(\lambda)$ is shown in Fig 4b for both vorticity production mechanisms—Coriolis and bottom frictional torques. An interval of length scale ratios centred around $\lambda \approx 1$ leads to a strong resonance in terms of residual tidal velocity. The latter falls off rather rapidly for both smaller and larger length scale ratios. This is because for very large topographic length scales vorticity gradients in the field vanish, so that no net flux of vorticity into a closed curve can be produced. For very

small length scales fluid columns experience during the tidal cycle such a rapid sign reversal of the vorticity production mechanisms that effectively no vorticity can be built up, so that no vorticity can be transferred to the mean field. The 'resonance peak' is reflected in the asymptotic behaviour¹⁵ of $g(\lambda)$

$$\begin{aligned} g(\lambda) &\propto \lambda, & \lambda \rightarrow 0 \\ g(\lambda) &\propto \lambda^{-1}, & \lambda \rightarrow \infty \end{aligned} \quad (12)$$

Hence, in view of the definition of λ ,

$$\begin{aligned} \tilde{u} &\propto U^2 \frac{h}{H} L^{-1}, & \lambda \rightarrow 0 \\ \tilde{u} &\propto \frac{h}{H} L, & \lambda \rightarrow \infty \end{aligned} \quad (13)$$

Thus, whereas the residual velocity increases with the square of the tidal velocity amplitude for large topographic length scales, it becomes independent of U for smaller length scales

Observations and model studies

Very few direct measurements of residual eddies exclusively from current velocity data exist, the evidence is usually indirect in that current velocity observations of eddy-like structures are complemented by knowledge about the bottom morphology or coastal geometry, or by modelling of the flow field. This particularly applies to the spatial structure of the eddies where, for complete spatial coverage in a rather small area, a very large array of instruments would have to be used simultaneously. The former difficulty can in principle be overcome by a new method of surface current measurements using Bragg scattering of HF-radar rays, used for the observation of surface wind waves. The surface current component in the radial direction from the antenna can be determined by observing the Doppler shift in the phase velocity of the waves. By using two antennas some distance apart the complete velocity vector can be measured. Hence an area of $\sim 2,500 \text{ km}^2$ can be mapped with a high resolution in time and space, allowing residual currents to be detected by averaging the observed velocity field in time over the dominant tidal period(s)¹⁷. At present, however, its accuracy, $\sim \pm 10 \text{ cm s}^{-1}$ for an individual current speed determination, means that the method is only successful for residual currents if the associated tidal currents exceed a value of 1 m s^{-1} , say, as the residuals are usually $< 10\%$ of the tidal currents proper.

Important parameter values of residual eddies are shown in Table 1, where the crucial parameter is the residual vorticity, which can vary over some orders of magnitude, from 10^{-4} to 10^{-7} s^{-1} , generally decreasing from the coastal zone to the open sea. It never seems to be larger than the so-called planetary vorticity (the Coriolis parameter), f , which at mid-latitudes is of the order of 10^{-4} s^{-1} . This may just be coincidence for basin or headland eddies (Fig 3a, b), but can easily be explained for sand-ridge eddies (Fig 3c, d) where the Coriolis force has an important role. In that case the residual vorticity is, in fact, derived from the planetary vorticity by vortex stretching and thus is unlikely to be larger than it. Typical associated velocities lie between 1 and 10 cm s^{-1} and length scales between 3 and 30 km .

In contrast to the eddy properties inferred from current measurements in the sea, results obtained exclusively from numerical or hydraulic modelling should be interpreted with care, as the appearance of eddies in these models depends critically on the basic equations, the boundary conditions and the grid size used in a numerical model and on the scaling used in a hydraulic model.

Analytical theories tend to use highly simplified geometries such as a rectangular basin⁹, a linearly sloping bottom¹³ or a bottom consisting of a superposition of plane horizontal waves^{14, 15, 18}. Although these approximations contain the basic physics, simulating the tidal velocity field, in particular the residual field, would enable us to cope with complicated situations that cannot be handled analytically. This can be done either by direct simulation in a hydraulic scale model or by

numerically solving the nonlinear equation of motion, (6), on a grid covering the investigated area, for some suitably chosen boundary conditions.

Successful hydraulic scale modelling depends critically on whether all dominant terms in the basic equation are scaled in the same proportion (similitude). If not, the flow field of the model is likely to be distorted compared with the field in the prototype. For residual eddies, only hydraulic models for the side-wall friction mechanism (Fig 3a, b), not incorporating rotational effects, have been described for an enclosed basin^{7, 19} and for eddies produced by headlands or islands in the Seto Sea (Japan)^{6, 20}. Similitude may be expected if the Reynolds number, R_e ,

$$R_e = \frac{UL_b}{\nu} \frac{L_b}{L} \quad (14)$$

based on the ratio of vorticity advection and vorticity diffusion, is the same in model and prototype, U being the tidal velocity amplitude, L the headland or basin dimension, ν the horizontal turbulent viscosity and L_b the thickness of the frictional boundary layer. Because $L_b \propto (\nu/\sigma)^{1/2}$, where σ is the tidal frequency, this leads to similitude of

$$R_e = \frac{U}{\sigma L} = \frac{\text{tidal excursion}}{\text{basin or headland length scale}} = \lambda \quad (15)$$

Again, the ratio between tidal excursion and topographic length scale is the basic dynamical parameter. The influence of the Earth's rotation makes hydraulic simulation much more complicated for the mechanisms displayed in Fig 3c, d because here, as in the prototype, the model should rotate. No results have yet been reported.

Numerical modelling results are only known^{8, 10-12, 21, 22} for situations which belong essentially to the mechanisms of Fig 3a, b. Side-wall friction is sometimes effectively produced by introducing a bottom sloping upwards to the coast, otherwise a no-slip condition at the boundary is applied. Anyhow, decreasing velocity towards the side-walls is crucial in this case because the absence of such velocity gradients means that no vorticity is present in the velocity field, so that no residual circulation can be built up. Along the boundaries these gradients are either

Table 2 Inventory of linear tidal sand ridges according to ref. 4, together with tidal current velocities and length scale ratios λ , assuming that tides are predominantly of the semidiurnal lunar type (M_2)

Area	Average ridge length (10^3 m)	Average tidal current velocity amplitude (m s^{-1})	λ
Great Bahama Bank	22.5	0.75	1.5
Rio Amazon	54.0	2.5	2.1
Sao Luiz do Maranhao (Brazil)	13.5	1.2	5.0
Malacca Strait	31.5	1.5	2.0
River Gambia to River Jeba (Africa)	15.0	0.75	1.3
Yalu River to Taech'ong Kudo	72.0	1.0	0.6
Wa Chiu Hsu to Tung-yin Shan	67.5	0.5	0.3
Keppel Isles to Whitsunday Island (Australia)	18.0	1.0	2.5
Prince of Wales Channel	13.5	0.5	1.7
Persian Gulf	13.5	1.0	3.3
	54.0	0.5	0.3
	54.0	1.0	0.3
	13.5	1.1	3.8
Southern Bight North Sea	49.5	0.7	0.6
Dover and Calais	27.0	1.1	1.8
	21.0	0.8	1.7
	21.0	0.8	1.7
	14.5	0.8	2.5
Rio Guayas Ecuador	13.5	1.1	3.3
Bahia Blanca Argentina	12.0	0.7	2.4
	12.0	0.5	1.9
Gulf of Maine/Georges Bank	36.0	1.0	1.3
Kuskokwim Bay, Alaska	27.0	0.9	1.5
	31.5	0.8	1.1
Nyshagak Bay, Alaska	18.0	1.2	2.9
Overall averages and standard deviations	29.0 ± 19	0.9 ± 0.4	1.9 ± 1.2

supplied by viscous friction at the wall or by differential bottom friction over a bottom sloping perpendicular to the wall. Hence a no-slip condition is essential when bottom slopes near the wall are absent or weak¹⁰. This condition is often not met in numerical modelling of tidal currents. Such models therefore do not produce residual eddies by the side-wall frictional mechanism, though they may be present in the prototype. Numerically it is not difficult to keep the Coriolis acceleration in the model, but its effect on eddies produced by side-wall friction seems to be minor¹⁰.

Finally, although there are no specific results from numerical modelling of the bottom slope mechanisms (Fig 3c, d), some of the residual eddies appearing in tidal models, such as those of the North Sea^{5,23}, should be ascribed to these mechanisms. However, the grid size of these models is generally much larger than the amplitude of the tidal excursion. Thus, the shape of the weighting function shown in Fig 4b suggests that the optimum topographic length scales for production of residual circulation are not resolved by such models. Residual circulation then only appears if the bottom topography fluctuates strongly on larger length scales and, in fact, it is in such a case that the models mentioned show residual circulation. Hence much of the fine structure in the residual field may be suppressed by choosing too large a grid size. This indicates that the curves in Fig 4b imply a criterion for the set-up of a suitable grid for the numerical calculation of residual circulation¹⁴, provided the bottom topography is properly reproduced on all length scales from the grid scale upwards. Residual circulation induced by the topography can be modelled successfully if the grid size is of the order of, for example, half the tidal displacement, in which case most of the responsive scales in terms of residual velocity are covered. A much smaller grid would be more time consuming in the computation without showing much more structure in the residual velocity field. On the other hand, a larger grid size does not resolve the optimum scales where eddies may be found.

Residual eddy diffusion

Tidal currents carry water parcels only to and fro or around in a closed orbit. No water parcel can escape these predetermined pathways unless some other velocity component allows a jump from one tidal pathway to another. If these jumps occur randomly, the rate of change of the spread of an ensemble of water parcels can be used to define an effective diffusion coefficient. Normally, in a turbulent fluid, such a coefficient can be expressed in the velocity and length scales of the largest turbulent eddies present in the flow. This, of course, also applies to turbulent tidal flow regimes, where residual eddies constitute the largest scales, though it is not *a priori* clear that such eddies diffuse matter in the same way as real turbulence does. Nevertheless, it is tempting to assume that the horizontal 'residual eddy diffusion coefficient', K_e , can be expressed as

$$K_e = c\tilde{u}L \quad (16)$$

where $\tilde{u}$ and L are characteristic residual velocity and length scales. This expression has little use unless the dimensionless constant, c , can be estimated. For two areas, the Seto Inland Sea in Japan⁶ and the Dutch Wadden Sea^{16,24,25}, this could be achieved by estimating K_e from the salinity budget of the area and, independently, by estimating characteristic values for u and L from the current velocity field. For the Seto Inland Sea c appeared to be of the order of 0.4–0.5, whereas for the Dutch Wadden Sea values between 0.5 and 2.0 were obtained, compatible with diffusion coefficients of the order of 1,500 m² s⁻¹ and 400–1,500 m² s⁻¹ respectively. Thus, if equation (16) is a correct expression for K_e , c seems not to be a real constant, but must depend on (an) additional parameter(s). A guess^{16,24,26} is that these parameters are the velocity and length scale ratios λ and ν encountered previously (equations (10) and (11)). Different values of c may then be explained by these ratios (particularly λ) varying between areas. Table 1 shows that representative ranges of ν and λ are 0.1–0.2 and 0.5–3.0 respectively. The few theories^{16,24,26} which explain the depen-

dence of c on ν and λ then give a range of 0.2–2.0 for c , which is of the same order as of the values derived empirically from the salinity budget in the case studies mentioned previously.

Geomorphological implications

The shape of the coastal geometry or the bottom morphology seems to be important to any tidal residual eddy theory, but is there also a reverse effect of the residual eddies on the morphology, does any specific shape of the morphology depend on the characteristics of residual eddy circulation? Such a connection is most obvious for eddies of the type shown in Fig 3c, d. The bottom morphological structures involved are the so-called 'linear tidal sand ridges', the worldwide appearance of which has been reviewed in ref. 4. Table 2 contains information about tidal velocity amplitudes for the areas discussed in ref. 4 together with estimates of the length scale ratio λ —the fundamental parameter in residual eddy dynamics. The average length of the sand ridges is 29 ± 18 km, whereas the average tidal velocity amplitude is 0.9 ± 0.4 m s⁻¹. More relevant is the average value and spread of λ , which reads 1.9 ± 1.2 . As Fig 4b shows this is exactly the range where maximum response of residual circulation can be expected. More recent studies^{27–29} of areas not included in ref. 4 show the same order of magnitude for the length of the ridges relative to the tidal excursion amplitude. The observation that the value of λ , on average, coincides with that for maximum response of residual circulation around the ridges, suggests that during their formation the ridges somehow adjust their shape so as to maintain the largest possible residual velocities around their contours. This suggestion is intensified by the observation^{29,30} that linear sand ridges are not exactly aligned along the major axis of the tidal current ellipse, but are offset in an anticlockwise direction (8–15°) as far as ridges on the Northern Hemisphere are concerned (no observations are at hand of ridges on the Southern Hemisphere but these must show the reverse effect, clockwise inclination). As shown in Fig 3c, d

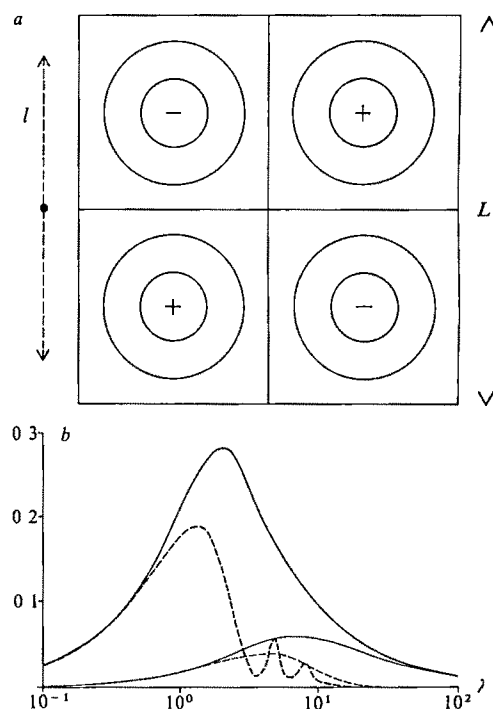


Fig 4 a, A test topography consisting of a regular series of humbs and bumps of height (+) or depth (-)h below the mean depth H . The wavelength L should be compared with the amplitude of tidal excursion l . The weighting function $g(\lambda)$ (refs 14, 15) for the Coriolis effect (solid curves) and for the bottom frictional effect (dashed curves), both for two values of the inverse bottom friction time scale r/H in equation (6), $r/H = 10^{-4}$ s⁻¹ (the highest solid and dashed curves) and $r/H = 10^{-3}$ s⁻¹ (the lowest solid and dashed curves). Note that for both vorticity production mechanisms $g(\lambda)$ has a maximum for λ in the range 1–5. The position of the maximum shifts to higher values of λ for increasing bottom friction coefficient or, equivalently, for decreasing mean water depth.

the ridges inclined anticlockwise with respect to the tidal current direction produce the largest response of residual circulation, as in that case bottom friction is reinforcing the Coriolis effect in setting-up a clockwise circulation around the bank. The geomorphological evidence shows, therefore, that both length scale and inclination of the ridges relative to the tidal current are generally such that maximum residual circulation around the ridge is produced, suggesting that this is necessary to maintain the ridge shape, which otherwise might degrade in the presence of other flattening or erosive processes. This may have some bearing on the interpretation of palaeo-environments, as these ridges are also observed in areas where a strong tidal influence no longer exists or where such ridges are buried under more recent sediments^{29,31}. For instance, the relict linear ridges found north-west of the Dogger Bank in the North Sea³¹ are thought to have been formed at ~7,000 BC when the sea level was 40–45 m lower. Whereas tidal currents are running north–south presently at the site, with a speed of $\sim 0.5 \text{ m s}^{-1}$, the orientation of the ridges implies that, at the time of their active formation, currents were predominantly running (north)east–(south)west, more or less parallel to a shoreline being formed by the northern flank of the Dogger Bank. According to the length of these relict ridges, on average 33 km (ref. 31), tidal current speeds must have been at least twice as high as they are there now.

In the case of headland eddies (Fig. 3b) it is less clear that the headland geometry itself is affected by the tidal current regime. In fact, nontidal factors principally determine the shape of the coastline. However, once the headland exists, it may produce a residual vortex pair. These vortices, in turn, collect sediment, thereby leading to the formation of sand banks^{11,12,32}. If centrifugal and Coriolis forces are both of equal importance in the eddies, sediment may accumulate asymmetrically between the eddies forming a vortex pair. The eddy showing anticlockwise circulation is preferred (on the Northern Hemisphere), because then Coriolis and centrifugal forces together with bottom friction reinforce each other in producing a convergence of bottom currents to the centre of the eddy¹¹. Thus, whereas the principal residual eddies do not depend on the existence of sand banks but on the headland geometry, imbedded within these principal eddies may be smaller ones coupled to the sand banks, which in turn are due to the sediment concentrating capacity of the principal eddies.

Finally, the geomorphological significance of the basin eddy in Fig. 3a is uncertain because it contradicts observations that in tidal channels the flood normally exceeds the ebb on the shallow sides of the channel, whereas the centre of the channel is ebb dominated^{33–36}. The sides, however, are often formed by drying tidal flats and there has been no explanation for the peculiar effects that may arise from areas uncovered by water for some time during the tidal cycle. The meandering of ebb-dominated gullies can set up a system of smaller residual cells within the overall residual pattern^{35,37,38}. It remains to be seen whether the dimensions of these cells are within the range of maximum

response for bottom morphological eddies, an observation from the Dutch Wadden Sea indicates that they may²⁴. In contrast to more open-sea areas the residual eddies suggested by ebb and flood channels in these shallow inland waters have a smaller dimension compared with the tidal excursion (a larger λ). This agrees with the result that a decrease of the bottom friction time scale, which is equivalent to a decrease in water depth, shifts the resonance peak to higher values of λ , as can be seen in Fig. 4b.

Conclusion

The existence of residual eddies in tidal areas can be largely explained by quasi two-dimensional vorticity dynamics together with a coastal geometry and/or bottom morphology providing tidal vorticity gradients. This theory does not really explain the characteristics of bottom morphological structures from the residual current velocity pattern, however. First, in residual sediment transport not only residual currents should be taken into account but also the nonlinearly generated higher harmonics in the tidal velocity field, the so-called over-tides^{12,39,40}. Yet, if sediment transport can be expressed as a function of the vertically averaged velocity, this may not be a serious flaw of quasi two-dimensional theory as over-tides appear as a by-product of residual current theory^{12–15}. Hence, a two-dimensional approach requires incorporation into the theory of a description of the interaction between (time-varying) sea-bed morphology and the tidal current regime. Such a theory might also give a first explanation of whether the bottom morphology is static or whether it may shift in space over a long time as is sometimes observed^{27,41}. However, the dimensionality of the present theories is a severe restriction. For problems of suspended sediment transport, the tidally varying vertical distribution of suspended load and velocity is important in producing residual transport, whereas for transport as bed-load the direction of the bottom current is decisive. In particular, the veering of the current direction vertically determines whether bottom currents, carrying the bulk of the sediment load, converge or diverge around any pre-existing bottom structure. Linear sand ridges, for instance, can only be maintained if there is a convergence of sediments towards the ridge crest, as is also observed indirectly by the inferred direction of bed-load transport from the properties of small sand waves on the sides of the ridge³⁰. Observational evidence on veering of tidal and residual currents near sand ridges is scarce, although the few observations we have^{42,43} support the conclusion derived from the sedimentary record³⁰. However, such observations cannot be interpreted in terms of our two-dimensional theories, and we lack the relevant three-dimensional theory.

I thank Drs J. H. Simpson and P. F. Linden, convenors of the Geophysical Fluid Dynamics Symposium at the 7th General Meeting of the European Geophysical Society, Budapest, 25–29 August 1980, who invited a review lecture on tidal residual circulation, of which the present paper is an extended version.

- 1 Van Veen, J. *Onderzoekingen in de Hoofden* (Algemeene Landsdrukkerij Den Haag 1936).
- 2 Van Veen, J. *Tijdschr. K. ned. Aardrijksk. Genoot* **54**, 155–195 (1937).
- 3 Van Veen, J. *J. Cons. perm. int. Explor. Mer* **13**, 7–36 (1938).
- 4 Off, Th. *Bull. Am. Ass. petrol. Geol.* **47**, 324–341 (1963).
- 5 Nihoul, J. C. J. & Ronday, F. C. *Tellus* **27**, 484–489 (1975).
- 6 Sugimoto, T. *J. oceanogr. Soc. Jap.* **31**, 1–14 (1975).
- 7 Yanagi, T. *J. oceanogr. Soc. Jap.* **32**, 199–208 (1976).
- 8 Oonishi, Y. *J. oceanogr. Soc. Jap.* **33**, 207–218 (1977).
- 9 Yasuda, H. *J. oceanogr. Soc. Jap.* **35**, 241–252 (1980).
- 10 Tee, K. T. *J. mar. Res.* **34**, 603–628 (1976).
- 11 Pingree, R. D. *J. mar. Biol. Ass. U.K.* **58**, 211–226 (1978).
- 12 Pingree, R. D. & Maddock, L. *Mar. Geol.* **32**, 269–289 (1979).
- 13 Huthnance, J. M. *Estuar. Coast. mar. Sci.* **1**, 89–99 (1973).
- 14 Zimmerman, J. T. F. *Geophys. Astrophys. Fluid Dyn.* **11**, 35–47 (1978).
- 15 Zimmerman, J. T. F. *J. mar. Res.* **38**, 601–630 (1980).
- 16 Zimmerman, J. T. F. in *Hydrodynamics of Estuaries and Fjords* (ed. Nihoul, J. C. J.) 207–216 (Elsevier, Amsterdam, 1978).
- 17 Frisch, A. S. & Weber, B. L. *J. geophys. Res.* **85**, 485–493 (1980).
- 18 Komen, G. J. & Rieppma, H. W. *Geophys. Astrophys. Fluid Dyn.* (submitted).
- 19 Yanagi, T. *J. oceanogr. Soc. Jap.* **34**, 67–72 (1978).
- 20 Higuchi, H. & Sugimoto, T. *Rapp. P.-V. Reun. Cons. perm. int. Expl. Mer* **167**, 177–184 (1974).
- 21 Pingree, R. D. & Maddock, L. *J. mar. Biol. Ass. U.K.* **57**, 869–875 (1977).
- 22 Pingree, R. D. & Maddock, L. *J. mar. Biol. Ass. U.K.* **57**, 339–354 (1977).
- 23 Prandle, D. *Proc. R. Soc. A* **359**, 189–228 (1978).
- 24 Zimmerman, J. T. F. *Neth. J. Sea Res.* **10**, 397–439 (1976).
- 25 Zimmerman, J. T. F. *Meth. J. Sea Res.* **10**, 149–191 (1976).
- 26 Oonishi, Y. & Kunishi, H. *J. oceanogr. Soc. Jap.* **35**, 136–140 (1979).
- 27 Caston, V. N. D. *Sedimentology* **18**, 63–78 (1972).
- 28 Nelson, C. H., Field, M. E. & Dupre, W. R. in *Abstr. int. Meet. Holocene Marine Sedimentation in the North Sea Basin* (eds Nio, S. D., Schuttenhelm, R. T. & Van Weering, T. C. E.) 90 (Texel, 1979).
- 29 Kenyon, N. H., Belderson, R. H., Stride, A. H. & Johnson, M. A. in *Abstr. int. Meet. Holocene Sedimentation in the North Sea Basin* (eds Nio, S. D., Schuttenhelm, R. T. & Van Weering, T. C. E.) 49 (Texel, 1980).
- 30 Caston, V. N. D. & Stride, A. H. *Mar. Geol.* **9**, M38–M42 (1970).
- 31 Jansen, J. H. F. *Neth. J. Sea Res.* **10**, 1–43 (1976).
- 32 Ferentinos, G. and Collins, M. *J. oceanogr. Soc. Jap.* **35**, 65–74 (1979).
- 33 Zimmerman, J. T. F. *Neth. J. Sea Res.* **8**, 126–138 (1974).
- 34 Charlton, J. A., McNicoll, W. & West, J. R. *Proc. R. Soc. Edinburgh B* **75**, 11–27 (1975).
- 35 Robinson, A. H. W. *Geography* **45**, 183–199 (1960).
- 36 Robinson, A. H. W. *Z. Geomorph. Suppl. Bnd 7*, 1–22 (1965).
- 37 Van Veen, J. *Tijdschr. K. ned. Aardrijksk. Genoot* **67**, 43–65 (1950).
- 38 Jacobsen, A. B. *Geogr. Tidsskr.* **61**, 119–141 (1962).
- 39 Groen, P. *Neth. J. Sea Res.* **3**, 564–574 (1967).
- 40 Pingree, R. D. & Griffiths, D. K. *J. mar. Biol. Ass. U.K.* **59**, 497–513 (1979).
- 41 Robinson, A. H. W. *Mar. Geol.* **4**, 57–84 (1966).
- 42 Harvey, J. G. & Vincent, C. E. *Estuar. Coast. mar. Sci.* **5**, 715–731 (1977).
- 43 McCave, I. N. *Mar. Geol.* **19**, 229–244 (1979).
- 44 Yanagi, T. *J. oceanogr. Soc. Jap.* **33**, 335–339 (1977).
- 45 Tee, K. T. *J. phys. Oceanogr.* **7**, 396–402 (1977).

ARTICLES

Photoelectron spectromicroscopy

G. Beamson, H. Q. Porter & D. W. Turner

Physical Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QZ, UK

Photoelectron spectroscopy is a powerful technique for studying electronic structure both of isolated molecules in the gas phase and of surfaces and adsorbed layers. Surfaces have also been imaged using photoemission microscopy. A new technique is described for combining the features of position imaging and energy analysis of the photoemitted electrons. The imaging is achieved using image projection in a divergent magnetic field to preserve the electron energy distribution.

DURING the construction of an instrument for the study of ions trapped in a high magnetic field we have explored other possibilities which this unusual arrangement of a magnetically collimated photoelectron energy analyser offers. In particular we considered imaging a surface using energy-analysed electrons. The redistribution between transverse and axial energy has a central role which depends on the divergence of the magnetic field, and this divergence produces magnification. Thus microscopy using energy selected photoelectrons is a possibility¹.

Energy analysing photoelectron microscopy

The new technique using photoelectron emission which may be seen as an extension of earlier experiments with conventional electron microscopes²⁻²², but introduces a new principle of electrostatic field-free imaging. The incident radiation (for

example, He(I) α) has a much higher energy (21.2 eV) than the work function. This enriches the image contrast by a wide range of photoelectrons from different valence levels. Energy analysis of the photoelectrons then allows spectroscopic analysis of the image.

The new technique depends on the capability of an axially symmetric divergent magnetic field to generate the enlarged image of a photoemissive object while preserving the original total energy distribution. Because the orbital moment of an electron moving in a divergent magnetic field is conserved as well as its total energy, any initially transverse energy reappears as energy of axial motion¹.

$$E = E_{\text{total}} - E_{\text{cyclotron}} \frac{B(2)}{B(1)}$$

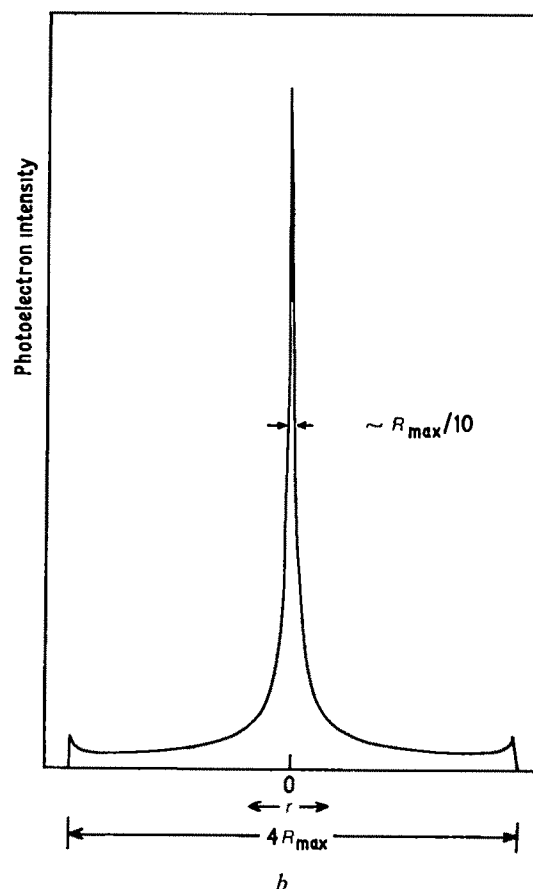
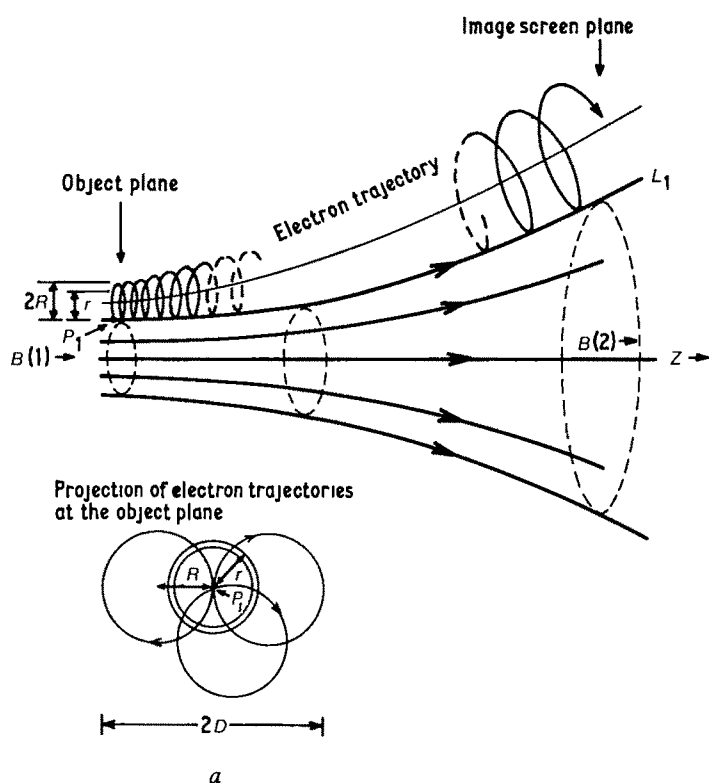


Fig. 1 *a*, Motion of a photoelectron in a divergent magnetic field. The electron is emitted at P_1 on the flux line L_1 . Inset, projection of trajectories at the object plane which determine a maximum circle of confusion (see text). *b*, The calculated intensity of photoelectron flux across the image of a point source for $\beta = 0$. The dimension ($4R_{\text{max}}$) is determined by the largest possible orbit radius R_{max} at the object surface for a given value of magnetic field and total electron energy (see text).

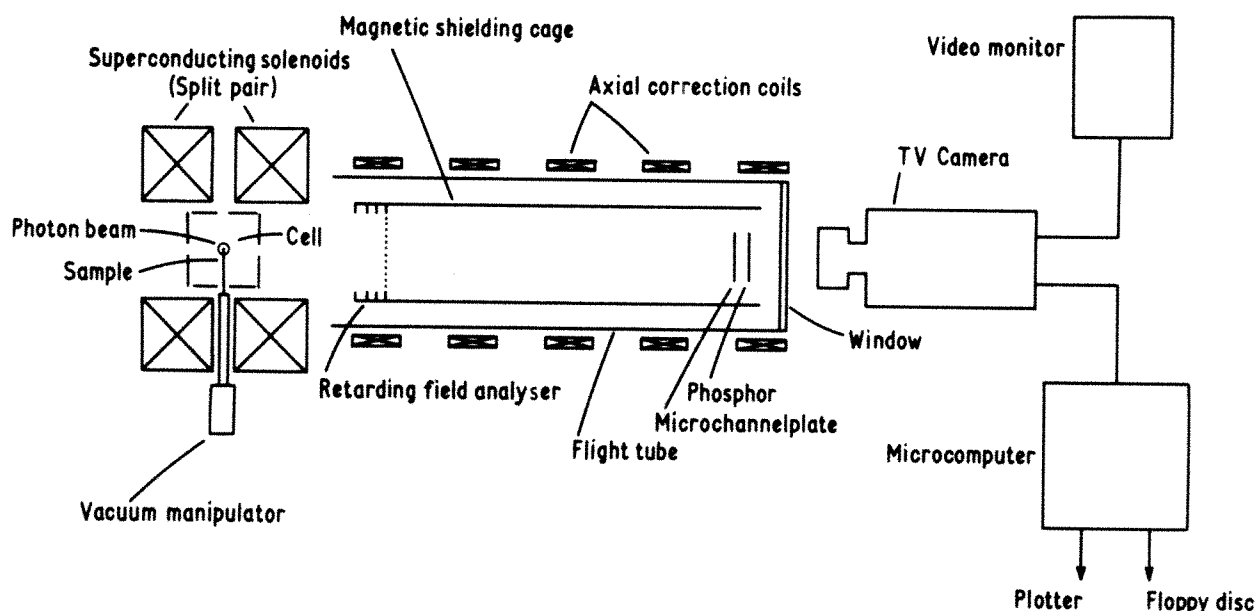


Fig. 2 The magnetic projection photoelectron microscope.

where E is the final energy of forward motion, E_{total} is the total electron energy, $E_{\text{cyclotron}}$ is the amount of energy initially in cyclotron motion and $B(2)/B(1)$ is the field ratio (Fig. 1a). As no focusing elements are used, this photoelectron microscope has the maximum possible luminosity. It is a magnetic analogue of Müller's "point projection electron microscope". It is, however, capable of bringing to an image screen every photoelectron liberated with a non-zero Z -component of velocity with exactly its original energy.

The important features of the motion of an electron which originates from an object point P_1 in a high magnetic field (Fig. 1a) are, first, that the helical motion about the adjacent flux line adjusts itself to conserve the orbital moment as the flux lines diverge, so that the orbit radius is proportional to the magnification, the area magnification is $B(1)/B(2)$ and the lateral magnification is $(B(1)/B(2))^{1/2}$. Then, second and as a consequence, after a sufficiently large change of field the original photoelectron energy is concentrated in the axial motion. The first inference implies that the electrons return repeatedly to the flux line passing through the object point P_1 ; the second, that planar field retarding potential energy analysis can be performed on the total photoelectron flux at any convenient image plane in a low magnetic field region.

Lateral resolution

The resolution of such a photoelectron microscope will be limited (apart from any limitation imposed by the method used to observe the image) by deviation of an electron from the guiding field line, which is in turn determined by the original angular distribution of photoemission and the kinetic energy of the photoelectrons emerging from the surface of the object. The electrons emitted otherwise than parallel to the magnetic field execute a cyclotron orbit about the field lines and their original point of emergence is uncertain by a maximum of twice the maximum orbit diameter (D) at the surface of the object (see Fig. 1a). The cyclotron radius of an electron of total energy E_{total} emitted at an angle θ to the magnetic field direction is given by

$$R = \frac{(2mE_{\text{total}})^{1/2} \sin \theta}{eB}$$

where B is the magnetic field strength. The probability of an electron of cyclotron radius R being found within a radial distance r from the field line passing through its point of origin (Fig. 1a, inset) is given by

$$\frac{2}{\pi} \sin^{-1} \left(\frac{r}{2R} \right)$$

This probability must be weighted according to the original angular distribution of photoemission referred to the z -axis²³. The radial intensity distribution can be calculated assuming an appropriate value of β , the angular distribution parameter. Figure 1b shows the radial intensity distribution plotted for $\beta = 0$. Changes in β have little effect on the general shape of this function, which is dominated by the fact that all electrons must return repeatedly to the original field line L_1 . This function is so strongly weighted towards small r values, that we predict an ultimate effective lateral resolution (the diameter of a 'circle of confusion') of approximately $R_{\text{max}}/10$ arising from the cyclotron motion, where R_{max} is $(2mE_{\text{total}})^{1/2}/eB$.

Table 1 shows this expected lateral resolution for some useful electron energy and central (strong) magnetic field values, with the de Broglie wavelength of the free electron (λ_e) for comparison.

It is useful to compare the present method with other electron microscopies that provide chemical information. In electron probe microanalysis and scanning Auger electron microscopy, the lateral resolution is limited by the area of the probe electron beam at the surface and is typically 0.3–0.5 μm using a thermal electron source. In scanning Auger microscopy with a cold-cathode field emitting source, the lateral resolution is limited²⁴ to ~30 nm and further improvement is prevented by electron scattering in the sample.

Thus from Table 1, in the present instrument, with a field strength of ~8 T, a useful microscopic performance is obtainable. For example, with 5-eV electrons, we could expect 0.1 μm resolution and in the highest readily attainable superconducting solenoid field, of ~16 T, performance should be at least comparable in lateral resolution to that of scanning Auger electron microscopy when lowest energy ($\approx kT$) photoelectrons are selected using samples at ordinary temperatures.

Depth resolution

Depth resolution (as opposed to lateral resolution) is sometimes referred to as depth of information and is related to the escape depth. (This should not be confused with depth of field, which is an instrumental property of a lens system and does not concern us because the present instrument is non-focusing; the field depth is limited only by the range of illumination of the object.) Thus a depth of information of d nm indicates that the final image is formed by photoelectrons that originate within a distance d of the surface.

The penetration depth of UV radiation is of the order of 1 μm but the electron escape depth is of the order of 0.5–20 nm. The precise value depends on the angle of exit and energy of the

Table 1 Lateral resolution of the magnetic projection photoelectron microscope for various electron energies and central magnetic fields

Electron energy (eV)	Expected lateral resolution (μm) for central magnetic fields (T)				de Broglie wavelength (μm)
	1	4	8	16	
500	8.0	2.0	1.0	0.5	5×10^{-5}
5	0.8	0.2	0.1	0.05	5×10^{-4}
0.05	0.08	0.02	0.01	5×10^{-3}	5×10^{-3}

photoelectron²⁵. The information depth is between 1 and 20 layers. The depth resolution in photoelectron microscopy is thus comparable with Auger electron microscopy and much higher than scanning electron microscopy. This high depth resolution means that in the study of metal surfaces for example, only the top few layers of atoms will contribute to the image. For biological surfaces, a depth of information of a few nanometres would permit a plasma membrane to be imaged while discriminating against the cytoplasm (see, for example, ref. 26). Depth of information is a function of incident light wavelength, angle of incidence and of any surface monolayers. Use of monolayer coating with a judicious choice of incident light parameters might further reduce the depth of information.

Energy resolution

By observing changes in the image as a function of the secondary electron energy distribution, scanning electron microscopy can, for example, discriminate between features whose secondary electron yield functions differ by as little as 0.5 eV. Con-

ventional UV photoelectron spectroscopy, however, has afforded energy discrimination as small as a few meV when isolated molecules are photoionized in the vapour phase²⁷ and, although band broadening effects in solids generally make the extremely high resolution of such experiments unnecessary, the possibilities for mapping a surface with such high energy resolution are extremely attractive. For example, sharp peaks (< 0.1 eV) have been seen in angle-resolved photoelectron spectra from surface states in copper²⁸ and gold²⁹. Some semiconductors and surface layers of weakly adsorbed species also exhibit narrow bands. Using the present instrument as a retarding potential analyser for gases gives an energy resolution of 20 meV—this may be capable of further improvement.

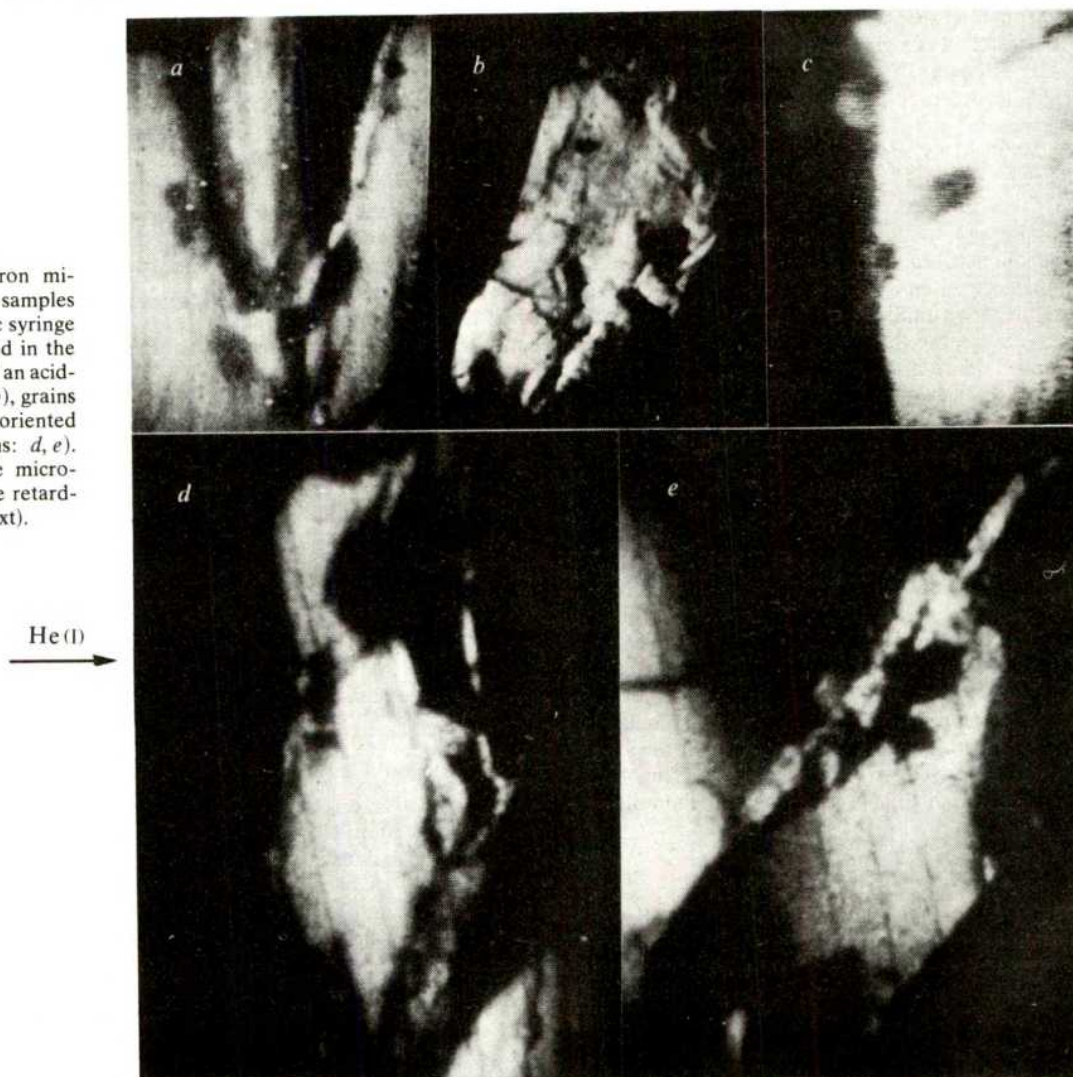
The present instrument

Figure 2 illustrates the present instrument. The divergent magnetic field is produced by a 'split-pair' superconducting solenoid whose central field can vary from 0 to 7.7 T. A gold-coated cubic electrode structure at the centre of the solenoid establishes a unipotential region, with various apertures for a 1-mm diameter collimated photon beam (initially of 21.2 eV photons from a He(I) d.c. discharge lamp), a sample probe, and for the emergent photoelectrons. The probe bearing the solid sample is mounted on the arm of a vacuum manipulator and can be translated and rotated.

Outside the solenoid, the magnetic field strength along the instrument axis is controlled using a magnetic shielding cage within the electron flight tube, and a system of axial and transverse correction coils outside the flight tube.

The magnetic shielding cage is an assembly of 24 gold-coated silver steel rods which reduce the magnetic field produced by the

Fig. 3 He(I) photoelectron micrographs of some test samples mounted on a hypodermic syringe needle. Brass wire inserted in the small hole (a). The edge of an acid-etched copper fragment (b), grains of pollen (c), a flake of oriented graphite (two orientations: d, e). Dark lines seen in these micrographs are shadows of the retarding grid wires (see text).



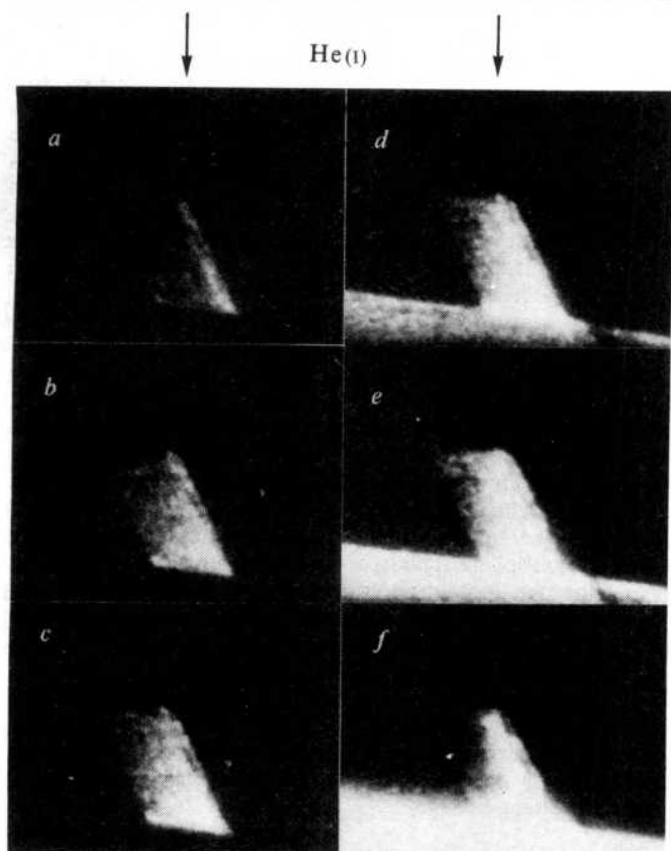


Fig. 4 He(I) photoelectron micrographs of a fragment of crystal-line chromium trisacetyl acetate attached to a stainless steel needle using electrons of energy greater than: a, 14 eV; b, 13 eV; c, 12 eV; d, 11 eV; e, 10 eV; f, 0 eV. Magnetic field strength 5 T. Thickness of the crystal $\sim 100 \mu\text{m}$.

superconducting solenoid within the cylinder. This, together with the external axial coils, is designed to establish a field which falls from a value of the order of 0.01 T at the nearest end to ~ 0.001 T at the furthest when the main field is at its maximum value. This large field reduction provides a magnification of the order of 100 and electron energy resolution is limited only by surface micropotentials (see Fig. 2 of ref. 1). With the present

arrangement, the argon photoelectron spectrum obtained in integral mode shows an energy resolution of ~ 20 meV. The resolving power of the retarding analyser is approximately constant in the 1–10 eV kinetic energy range.

A secondary function of the gold-plated silver steel rods is to establish a uniform transverse electrostatic field as an alternative way of energy analysis. Mounted at the front end of the shielding cage is a gold-coated tungsten wire grid with 95% transparency which, together with a system of gold-coated stainless steel rings, establishes a planar retarding field energy analyser. At the outer end of the shielding cage is the electron detection and imaging system. This consists of a 2-inch diameter microchannel plate (25- μm channel spacing) and low energy phosphor screen.

The lateral resolution of object detail is limited to $\sim 2 \mu\text{m}$ mainly by the coarseness of the 'home-made' phosphor screen and vibrations at the object. Ultimately, the limit will be imposed by the channel plate structure and the finest channel plate structure available (12 μm) would correspond to a lateral resolution of the object of 0.12 μm when $B(1)/B(2)$ is 10^4 . The photographs in Figs 3–7 are degraded by marked 'blooming' in the TV image close to features having high intensity. Both of these defects should be removed by using a charge-coupled electron image sensor.

The He(I) photoelectron count rate at the phosphor screen is typically 50,000 c.p.s. from argon at 2×10^{-5} torr. With a solid sample, the electron flux is too large for electron counting. The image on the phosphor screen is observed by a television camera and can be photographed directly from a video monitor. Alternatively, the camera output can be digitized and stored in a micro-computer, either as the whole picture frame or just one line or section of a line. In the latter case, the computer is used to advance the retarding potential on the analyser as successive frames are scanned and so to build up an array of spectra along a particular picture line segment. This array of integral spectra can be processed to yield the more familiar form of differential photoelectron spectra (see, for example, Fig. 7).

The vacuum chamber of the instrument is pumped by a 5-inch diffusion pump through a water-cooled baffle and liquid nitrogen trap. Thus a background pressure of 10^{-7} torr is readily achieved when the helium lamp is operating.

Although the present instrument has no vacuum lock for sample changing nor a separate argon ion gun for surface etching, *in-situ* cleaning is straightforward, with the strong magnetic field playing an important part. When argon is admitted to the working region to a total pressure of $\sim 100 \mu\text{m}$, a

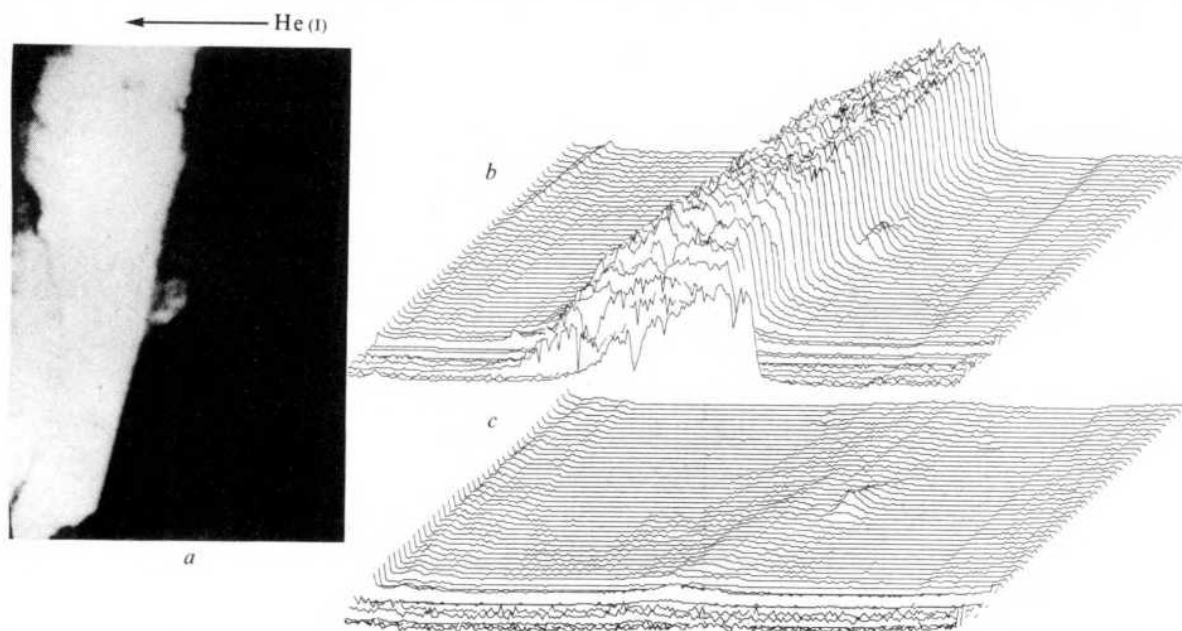


Fig. 5 a, Helium (I) photoelectron micrograph of pollen grains attached to a syringe needle. b, c, Digitally accumulated pictures of the same subject using electrons of energy greater than 5 eV (b) and 13.5 eV (c).

Penning-type glow discharge is established with ~ 400 V between the sample support and the surrounding cage. A current of ~ 0.1 mA for a few minutes causes rapid changes in the images of both metallic and plant material surfaces.

Examples

As a preliminary demonstration, a few fragments of mineral and biological origin were attached to the top of a hypodermic syringe needle placed on the manipulator. Figures 3–7 show images obtained by photographing the normal TV image on a monitor screen together with their electron energy distributions along a chosen TV raster line. The direction of illumination by the He(I) beam is indicated.

Figure 3 shows a haphazard collection of objects imaged using electrons of all energies, that is the retarding potential of the analyser was set to zero. The images are (a–e) of a 250 μm brass wire, the edge of an acid-etched annealed copper foil fragment, some pollen grains and two views of a flake of 'oriented graphite' (UCAR) stuck tangentially to the needle surface.

Images of a microcrystal of chromium (III) triacetylacetonate ($\text{Cr}(\text{Acac})_3$) some 100 μm thick obtained using energy analysed electrons are shown in Fig. 4. The image of the supporting surface (stainless steel) could be completely suppressed by observing only high energy electrons (> 11 eV). Much more significant is the variation of relative brightness of the different crystal faces at different voltages. One face in particular (the right-hand edge in Fig. 4), which is both illuminated and observed at near grazing incidence, remains visible in photoelectrons of energy 17 eV. At low electron energy the end of the crystal facing the beam is the strongest emitter (this is not clearly shown in the present series of photographs because the vidicon was overloaded).

Fragments of plant material (pollen grains, cellular fragments and fibres) showed similar behaviour when observed immediately after sample insertion and initial evacuation, or after brief argon ion bombardment. But the maximum kinetic energy for photoelectrons was generally lower (~ 14 eV) although higher energy photoelectrons could be observed from selected areas with carefully controlled ion-cleaning. Figure 5a shows the > 10 eV photoelectron image of a pair of pollen grains (diameter ~ 50 μm) on the stainless steel surface together with computer-stored images obtained using > 5 eV (Fig. 5b) and > 13.5 eV (Fig. 5c), photoelectrons. The pollen grains (of *Echinopsis eyriesii*) have slightly dimpled spherical surfaces but the images show a marked spatial dependence of the energy distribution.

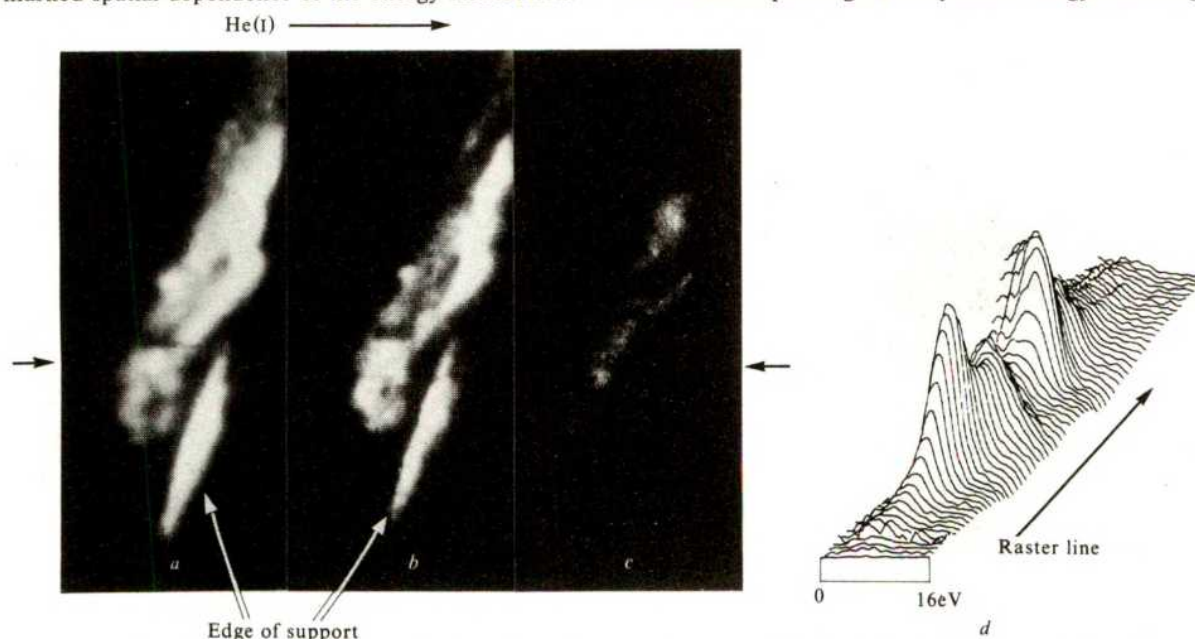


Fig. 6 a, b, c, He(I) photoelectron micrographs of a fragment of poppy anther using: a, all electrons; b, electrons of energy > 8 eV and c, electrons of energy > 11 eV. Magnetic field strength, 4 T. d, Photoelectron energy spectra (16 eV energy scale) along the raster line by arrows.

In a more complicated plant fragment (of poppy anther), an interesting variation of electron energy distribution along a selected line can be discerned (Fig. 6d) related to the intensity distribution in the different images obtained as the retarding potential varies (Fig. 6a, b, c).

Figure 7 shows the change in photoelectron emission from an acid-etched copper surface as it is cleaned by ion bombardment. Figure 7a is an all-photoelectron micrograph of the copper surface before cleaning. Figure 7b–d shows the changes which occur in the electron energy distribution along the arrowed line in Fig. 7a as the surface becomes cleaner. Figure 7b shows the uncleaned surface. Figure 7c, d shows the progressive changes after Ar^+ bombardment. A low-electron energy band is progressively lost. The final observed spectra do not have the appearance of conventional photoelectron spectra of copper surfaces for several reasons. First, although the annealing and acid etching did afford a surface topography showing the presence of different crystallites, the exposed faces were ill-defined crystallographically. Second, the distribution is not angle resolved. In effect we have integrated the spectra over the emitting hemisphere. This is, of course, a feature of the magnetic projection. All emitted photoelectrons are collected. Third, in this preliminary study the surface is not completely clean even after ion bombardment, in particular the Fermi edge is still suppressed.

The bright line on the left of Fig. 7a is photoemission from the epoxy resin used to mount the copper sample, together with its shadow. The difference in contrast arising from the different photoelectron energy distribution is quite striking at certain analyser voltages.

Conclusions

We have demonstrated that photoelectron images of microscopic surface structure can be obtained using energy-analysed electrons. The He(I) resonance line was used for the demonstration because it is easy to generate and has been widely used in the study of surface valence and photoemission. For element mapping, however, other resonance lines could be used. For core electrons, in particular, soft X-ray lines and synchrotron radiation may have advantages where it is possible to match the photon energy to a suitable shell or sub-shell ionization energy of the required element. This technique thus has the unique ability of producing either a molecular or atomic map of the surface depending on the photon energy. Scanning Auger

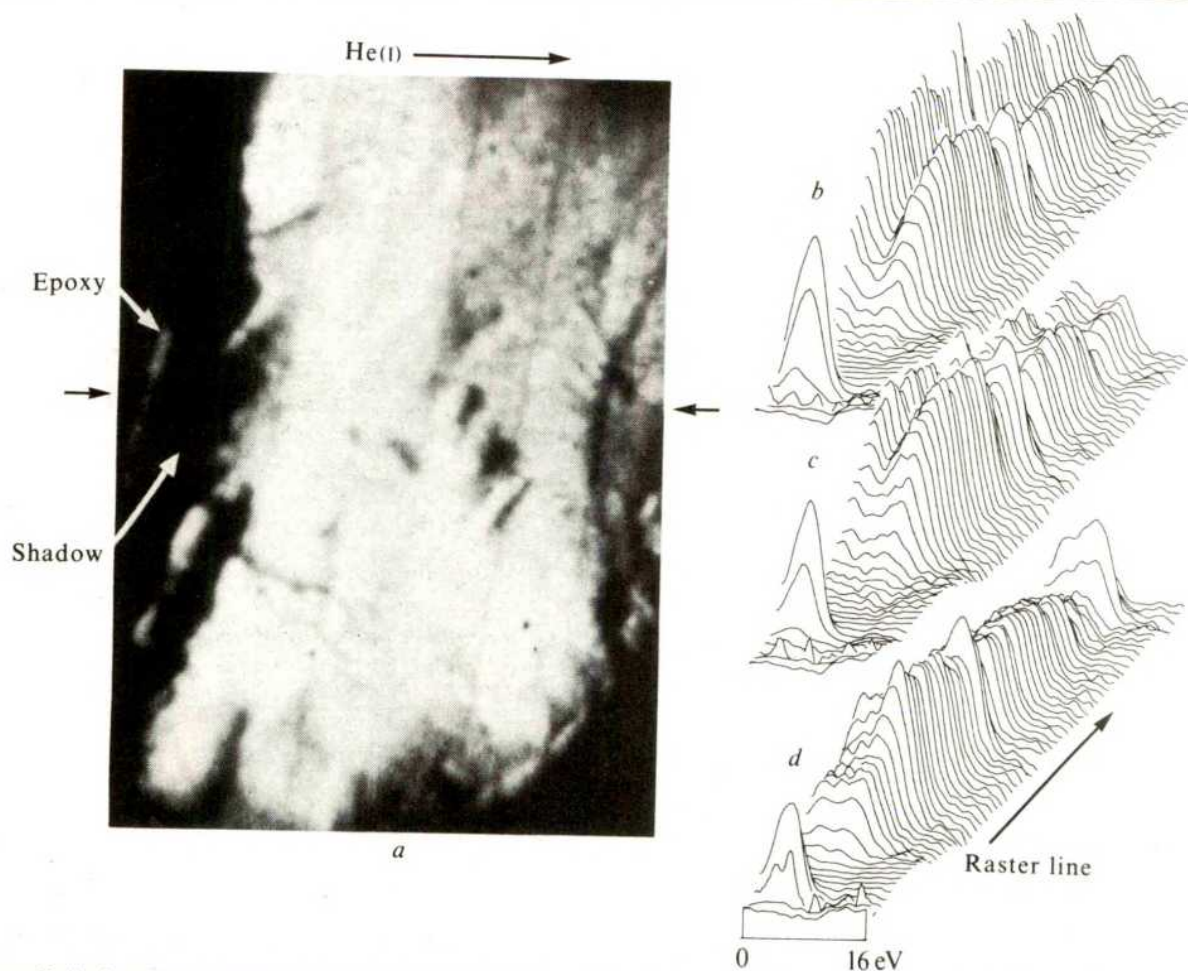


Fig. 7 *a*, He(I) photoelectron micrograph of an acid-etched copper surface before ion bombardment. Magnetic field strength, 7 T. *b*, *c*, *d*, The changes in photoelectron energy spectra along the raster line marked by arrows, as the surface is progressively cleaned by argon ion bombardment.

microscopy, on the other hand, provides only an atomic or elemental map of the surface. By using a photon energy only slightly in excess of the ionization energy both lateral resolution and sensitivity will be at the optimum. When the sample is irradiated near grazing incidence, photoemission is concentrated in the first atom layer and depth resolution is then at its optimum.

A most important advantage may, however, be derived from the much less destructive nature of the photoemission process combined with the high electron collecting efficiency.

To obtain good lateral resolution in Auger microscopy a current density of 10 A cm^{-2} is required²⁴. In such irradiation electron trapping, charging and secondary electron emission are important and the following electron-induced artefacts can be observed: migration of surface ions, decomposition, electron-induced desorption and diffusion induced by local Joule heating. The power dissipated in the sample by the He(I) discharge lamp is estimated to be less than that in the Auger electron micro-

scope by a factor of 1,000 or more. The advantage may be considerably larger, when the reduction in chemical damage associated with photons compared with electrons is taken into account.

The use of a very strong magnetic field, though primarily instrumental in intent, also has important possibilities for the study of energy bands and spin polarization in magnetic materials. This, when combined with the imaging property, may be particularly valuable in materials science in the study of magnetic domains. Photoemission from a ferromagnetic material in which the domains have been aligned by an external 0.05 T field is known to lead to spin-polarized photoelectrons³⁰.

This work was supported by the Paul Instrument Fund of the Royal Society, the SRC and the Nuffield Foundation. We thank G. I. Bird, H. Halleybone, I. R. Plummer for help in preparing the TV camera and programs, the PCL laboratory staff, particularly I. March and R. Arthur for assistance and S. Davies for advice.

Received 25 July 1980; accepted 4 February 1981.

1. Beamson G., Porter H. Q. & Turner D. W. *J. Phys. E: Scient. Instrum.* **13**, 64-66 (1980).
2. Goldstein, E. *Ann. Phys.* **11**, 832-856 (1880).
3. Knoll, M. & Ruska, E. *Ann. Phys.* **12**, 607-640 (1932); **12**, 641-661 (1932).
4. Knoll, M., Houtermans, F. G. & Schulze, W. *Z. Phys.* **78**, 340-362 (1932).
5. Brüche, E. & Johansson, H. *Ann. Phys.* **15**, 145-166 (1932).
6. Brüche, E. *Z. Phys.* **86**, 448-450 (1933).
7. Pohl, J. *J. tech. Phys.* **15**, 579-581 (1934).
8. Mahl, H. & Pohl, J. *J. tech. Phys.* **16**, 219-221 (1935).
9. Marton, L. *Physica* **3**, 959-967 (1936).
10. Müller, E. W. *Phys. Z.* **37**, 838-842 (1936).
11. Müller, E. W. *Z. Phys.* **108**, 668-680 (1938).
12. Kinder, E. *Naturwissenschaften* **30**, 591-592 (1942).
13. Huguenin, E. L. *Ann. Phys., Paris* **2**, 214-266 (1957).
14. Koch, W. *Z. Phys.* **152**, 1-18 (1958).
15. Engel, W. *6th int. Congr. Electron Microscopy*, Kyoto, 217 (1966).
16. Wegmann, L. *Prakt. Metallogr.* **5**, 241 (1968).
17. Möllenstedt, G. & Lenz, F. *Adv. Electron. electron Phys.* **18**, 251-283 (1963).

18. Wegmann, L. *J. Microsc.* **96**, 1-23 (1972).
19. Scott, G. H. & Packer, D. M. *Anat. Rec.* **74**, 17-29 (1939).
20. Griffith, O. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **69**, 561-565 (1972).
21. Houle, W. A., Brown, H. M. & Griffith, O. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4180-4184 (1979).
22. Engel, W. & Grund, S. *8th int. Congr. Electron Microscopy* Canberra, 656 (1974).
23. Beamson, G., Pearce, S. J. & Turner, D. W. *Phys. Scr.* **16**, 186-190 (1977).
24. Hofmann, S. in *Comprehensive Analytical Chemistry* Vol. 9 (ed. Svehla, G.) Elsevier, Amsterdam, 1979.
25. Brundle, C. R. *Surf. Sci.* **48**, 99 (1975).
26. Burke, C. A., Birrell, G. B., Lesch, G. M. & Griffith, O. H. *Photochem. Photobiol.* **19**, 29-34 (1974).
27. Turner, D. W., Baker, A. D., Baker, C. & Brundle, C. R. *Molecular Photoelectron Spectroscopy* (Wiley, London, 1970).
28. Heimann, P., Hermanson, J., Miosga, H. & Nedermeyer, H. *Phys. Rev. Lett.* **42**, 1782-1784 (1978); *Phys. Rev.* **B20**, 3059-3066 (1979).
29. Heimann, P., Hermanson, J., Miosga, H. & Nedermeyer, H. *Phys. Rev. Lett.* **43**, 1757-1760 (1979).
30. Pierce, D. T., Kuyatt, C. E. & Celotta, R. J. *Rev. Sci. Instrum.* **50**, 1467-1473 (1979).

Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes

Hitoshi Sakano, Yoshikazu Kurosawa, Martin Weigert & Susumu Tonegawa

Basel Institute for Immunology, 487 Grenzacherstrasse, Postfach, 4005 Basel 5, Switzerland

A putative diversity segment of immunoglobulin heavy-chain genes (D segment) has been identified 700 base pairs 5' to J_H DNA on the germ-line genome of the mouse. This 10-base pair D segment is flanked by two sets of sequences related to

CACTGTG and GGTTTTGT
GTGACAC and CCAAAAACA

which are possible recognition sites for a recombinase. The spacer separating the heptamer and the nonamer is 12 base pairs long on both sides of the D segment. As the spacer separating the two signal sequences in V_H DNAs and J_H DNAs is 23 ± 1 base pairs long, the two recombinations required for creation of a complete immunoglobulin V_H gene, a V_H -D joining and a D- J_H joining, follow a 12/23-base pair spacer rule. Allelic exclusion is discussed with respect to D segments.

It is well established that complete immunoglobulin variable region genes (V genes) are generated by somatic DNA recombination(s) during the development of B-lymphocyte cells¹⁻¹⁵. In the light-chain genes of the mouse, in both λ - and κ -types, the last 13 COOH-terminal residues of the variable region are encoded by a separate DNA segment called J DNA (J for joining) in the germ-line genome^{4,7,9,10}. Coding information for the rest of the V region is contained in an embryonic V DNA segment^{3,6,7,9,10}. Accordingly, somatic recombination occurs between the 3' end of the V DNA and the 5' end of the J DNA to generate a complete light-chain V gene^{4,6,8}. The $J\lambda_1$ (ref. 6) and five $J\kappa$ DNAs^{9,10} carry two blocks of sequences, one a palindromic heptamer, CACTGTG, and the other a nonamer, GGTTTTGT, highly conserved in the 5'-noncoding regions. These conserved sequences are invertedly repeated in the 3'-noncoding regions of all embryonic $V\lambda$ and $V\kappa$ DNAs sequenced to date^{3,6,7,9,10}. Another striking feature around the recombination site is that the spacer between the heptamer and the nonamer^{12,13} in all V and J DNAs is 12 ± 1 or 23 ± 1 base pairs long. In addition, recombination seems to occur only between a V DNA with the shorter (12-base pair) spacer and a J DNA with the longer (23-base pair) spacer or vice versa (12/23-base pair spacer rule). Based on these observations

Early *et al.*¹² and Sakano *et al.*¹³ predicted that the same or a similar enzyme mediates the recombinations in both λ - and κ -light-chain genes and that the recombinase consists of two functionally distinct units, one recognizing the heptamer and the nonamer separated by the shorter (12-base pair) spacer and the other recognizing the same signal sequences separated by the longer (23-base pair) spacer.

In contrast to the light-chain genes, heavy-chain J DNAs (J_H) do not seem to associate directly with the heavy-chain V DNAs (V_H). In the case of the $\gamma 2b$ heavy-chain gene expressed in myeloma MOPC141, the 14-residue peptide comprising the third hypervariable region (HV3) beginning with Val (residue 98) and ending with Thr (residue 111) is encoded neither in the germ-line V DNA nor in the J DNA¹³. A similar observation has been made independently in an α heavy-chain gene¹². Therefore, it is generally believed that the HV3 region of a heavy chain is encoded by a separate DNA segment, D (D segment for diversity¹⁶), and that two recombination events, V_H -D and D- J_H joinings, are necessary to form a complete heavy-chain V gene. All germ-line V_H DNAs and J_H DNAs sequenced so far have the conserved heptamer and nonamer separated by the longer spacer. It has therefore been predicted^{12,13} that germ-line D segments must contain the same recognition sequences on each side of the coding segment and their spacers must both be short ones, so that the 12/23-base pair spacer rule is adhered to. However, no germ-line D sequence has been reported to date and the predictions outlined above still need to be verified. It was previously observed that the DNA sequences are often rearranged in the J_H cluster in some cloned T-cell lines^{17,18}, as well as in pre-B-cell lines^{18,19}. By the Southern hybridization method using a J_H cluster probe, we observed that DNA rearrangements within the J_H cluster region are not restricted to one allelic chromosome in most of the myelomas. One intriguing possibility was that some of these rearrangements were abortive, representing recombination between a D segment and a J_H segment producing an incompletely rearranged D- J_H structure. We therefore cloned some of the genomic DNA fragments containing the rearranged J_H -cluster sequence from myelomas and characterized the cloned DNA fragments. These studies led to an identification of a potential germ-line D DNA segment, which is reported here. This potential D DNA segment is flanked by the two pairs of conserved sequences in inverted orientations, and the spacers separating the two signal sequences are 12 base pairs long for both.

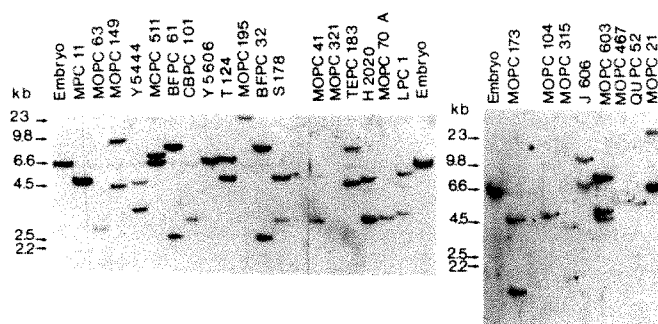
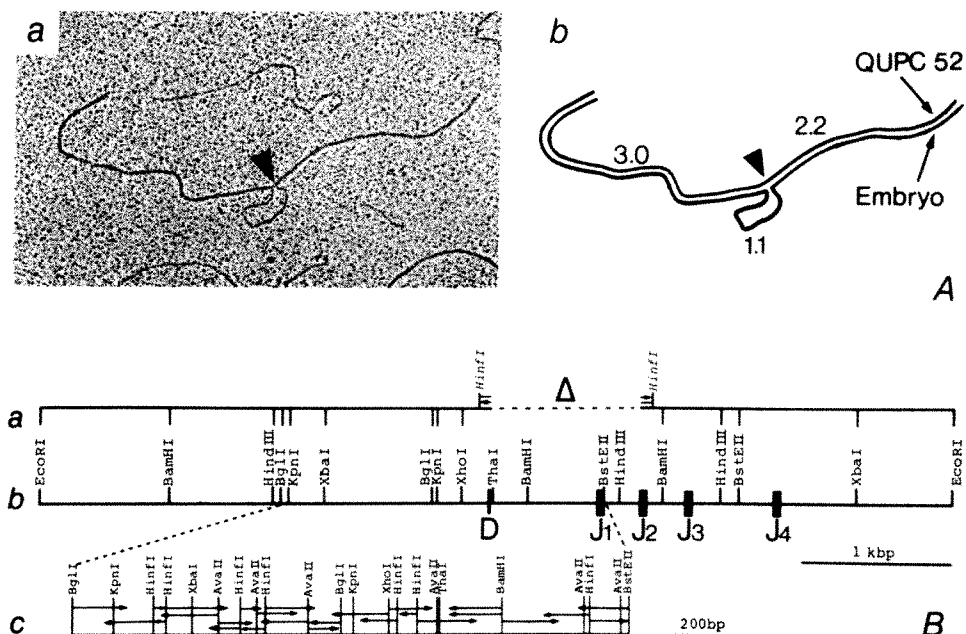


Fig. 1 Southern blot analysis of mouse embryo and myeloma DNAs. Mouse DNAs were digested with *Eco*RI, separated in 0.8% agarose gels and blotted to nitrocellulose filters according to the method of Southern²⁰. Filters were then incubated with the nick-translated J_H probe (1.1-kilobase *Sac*I-*Eco*RI fragment, see ref. 11), essentially as described by Wahl *et al.*³⁹. *Hind*III fragments of phage λ DNA were used as size markers (in kilobases, kb). For the cloning of a 5.2-kilobase Q52 fragment hybridizable with the J_H probe, *Eco*RI digest of the myeloma DNA was fractionated in a 0.8% agarose gel and DNA from the J_H probe-positive fraction was eluted as described previously⁴⁰. Recombinant phages containing the mouse DNA were screened by *in situ* plaque hybridization⁴¹ with the J_H probe.

Fig. 2 A, Electron micrograph (a) and schematic interpretation (b) of a heteroduplex molecule formed between two *Eco*RI inserts, one from clone ME184-8 (embryonic J) and the other from clone Q52J (rearranged J from myeloma QUPC52). Procedures used for the heteroduplex formation have been described previously⁴. Clone ME184-8 was prepared by Richard Maki and contains, in a phage λ WES vector, the 6.4-kilobase *Eco*RI embryo DNA fragment carrying the four J_H DNA segments whose sequences were determined previously¹³ using another DNA clone referred to as MEP203 (ref. 11). Numbers indicate lengths of various parts of the heteroduplex in kilobases. B, Restriction enzyme cleavage maps of the J-containing *Eco*RI fragments. a, 5.2-kilobase insert from the myeloma clone Q52J; b, 6.4-kilobase insert from the embryonic clone ME184-8. The broken line (Δ) in Ba represents a deletion in the myeloma clone; the bars in Bb indicate coding regions. The positions of four J_H segments have been determined by R-loop mapping and DNA sequencing^{11,13}. The D segment is identified in the present study by DNA sequencing of the 5' flanking region of J_{H1} . Strategy of the sequencing is shown in Bc. kbp, Kilobase pairs; bp, base pairs. The previously published *Bam*HI sites¹³ have been corrected.



Frequent rearrangements of the J_H region sequences in myelomas

We analysed *Eco*RI digests of total DNA from 18 different BALB/c myelomas by the Southern gel blotting method²⁰ using a J_H probe (Fig. 1). The embryo DNA gave a band of 6.4 kilobases as expected from previous studies¹¹. All but three (MOPC511, MOPC21 and MOPC321) myelomas gave one, two or three bands that are all different from the embryo band. MOPC511 and MOPC21 each gave two bands, one of which is indistinguishable from the embryo band. MOPC321 gave no band. This myeloma synthesizes no heavy chain and presumably lacks part of the heavy-chain gene sequences in question. In some myelomas a very faint band can be seen at the position of the embryo band, probably due to contamination of non-myeloma cells in the solid tumours from which the DNA was prepared. The variability in the number of bands among various myelomas and in the intensity of bands within a myeloma most probably reflects the fact that myelomas are usually not diploid²¹.

We assume that one rearranged band in each myeloma represents the V_H -D- J_H joining required for the formation of a complete heavy-chain gene active in that myeloma. Although the nature of the additional rearrangements observed in many myelomas is unknown, the results show that sequence rearrangements are very frequent in the vicinity of the J_H cluster in all copies of chromosome 12, on which the heavy-chain genes reside²².

DNA deletion on an abortively-rearranged J_H fragment in myeloma QUPC52

We chose the 5.2-kilobase QUPC52 (Q52) fragment for further analysis and cloned it in the phage vector λ WES (ref. 23). The cloned *Eco*RI fragment was analysed by heteroduplex formation using a germ-line J_H fragment (clone ME184-8) cloned from the *Eco*RI digest of mouse embryo DNA. As shown in Fig. 2A, the 5.2-kilobase myeloma fragment carries a 1.1-kilobase deletion, which lies 3.0–4.1 kilobases from the 5' end of the embryonic J_H fragment. By superimposing the heteroduplex map on the restriction enzyme map of the embryonic clone, we could locate the right end of the deletion near the J_{H2} DNA and the left end about 0.7 kilobases 5' to the J_{H1} DNA (Fig. 2B).

D-like segment is attached to J_H DNA on the abortive clone

To analyse the structure of the DNA segment joined to the J_H region, we purified a 600-base pair *Bgl*II-*Bam*HI fragment (see Fig. 2Ba) from the myeloma clone and determined the nucleotide sequence. Figure 3b shows the nucleotide sequence of a 110-base pair *Hinf*I-*Hinf*I fragment containing the deletion. As expected, we found the J_{H2} sequence in the sequenced region. By comparing this myeloma sequence with the germ-line J_{H2} sequence (Fig. 3b, c) we found that the J_{H2} sequence starts with the third nucleotide of the Asp codon at position 101 (number-

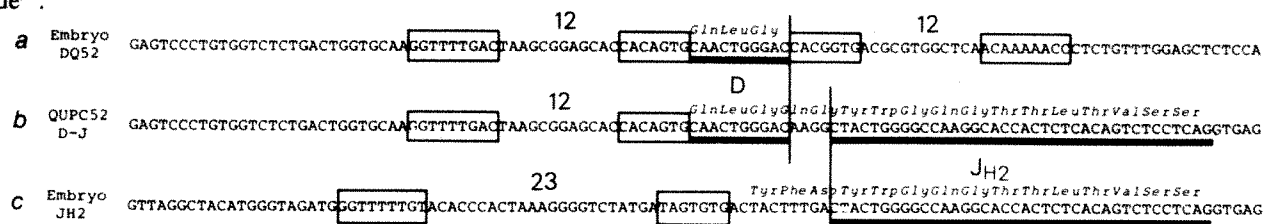


Fig. 3 Nucleotide sequence of the D-J structure in a myeloma clone, Q52J (b), and its germ-line sequences, J_{H2} (c) and D_{Q52} (a). For the sequencing of b, a 110-base pair *Hinf*I-*Hinf*I fragment containing a deletion was purified from a 600-base pair *Bgl*II-*Bam*HI fragment (see Fig. 1Ba), both ends labelled with [α -³²P]-nucleoside triphosphates, and the strand separated and sequenced in both directions according to the procedures described previously²². The embryonic J_{H2} sequence (c) was taken from ref. 13. The embryonic D_{Q52} region (a) was sequenced according to the strategy shown in Fig. 2Bc. Two blocks of conserved sequences,

GGTTTTTGT and CACTGTG
CCAAAACA and GTGACAC

or their related sequences are boxed. Numbers indicate lengths of spacers separating the two conserved sequences in base pairs. Sequences in a and c contributing the coding region in b are underlined. The corresponding sequences in b are also underlined. Note that tetranucleotide AAGG in b between the two boundaries of D_{Q52} and J_{H2} is not accounted for by either germ-line sequence. Vertical lines indicate possible recombination sites with D_{Q52} and J_{H2} . Encoded amino acid sequences are in italics.

KpnI
GGTACCCCCCATGGCTTCCCCGGAGGCACTGCCCAACCCCCACCATCACAGACCTTTCTCCATAGTTGATAGCTCAGCAGAGTGAATGACAGATGGACCTCCATCTGCTCTTATTTTAAAAA 120

XhoI
GAAGACAAACCCACAGGCTCGAGAACTTTAGCGACTGTTTGTGAGAGAAATCATTTGGTCCCTGACTCAAGAGATGACTGGCAGTTTGGGGATCAAGAATACCCATACTCTGTGGCTAGTG 240

12 D052 12
TGAGGTTTAAGCCTCAGAGTCCCTGTGGTCTCTGACTGGTGCAAGGTTTGGACTTAAGCGGAGCACACAGTGCAACTGGGACACGGTGACGCGTGGCTCAACAAAAACCTCTGTTTGG 360

AGCTCTCCAGGGGAGCCTGAGCTATGAGGAAGTAGAGAGGCTTGAGAAATCTGAGGAAGAAAAGAGTAGATCTGAGAGGAAAGGTAGCTTTCTGGAGGTGAGGAGACAGTGACAGAGAA 480

BamHI
AACGAGTTACTGTGGACAGGTCTTAGATGGGGAAAGAATGAGCAAAATGCAAGCATCAGAAGGATAGATGCAATGTCTTGCCAAAGGACTTACCAAGAGGATCCCGGACAGAGCAGGCAGGT 600

GGAGTTGACTGAGAGGACAGGGTAGGTGCAGGTTCCTCTCTCGTTTCCTTTCTCCTTCTCTGTTTCTCTCTCTCTGTGTGTCACAGGTCTCACTATGCTAGCCAAGGCTAGCCTGAAAGAT 720

TACCATCTACAGATGGGCCCATCCAGTTGAGTTAAGGTGGAAGATCTCTCCAAACATCTGAGTTTCTGAGGCTTGGATGCCACTGGGGACGCCAAGGAGCTTTGGGCTGGGTTGGTT 840

22 JH1 BstEII
GGCCCCACATGTAGGGCTACTTCACTGGGTCTATAATTAATCTGATGTGCTAGGACCAGGGGGCTCAGGTCACTCAGGTGAGTCTGATCTGGGGACTGTGGGGTTTCAGGTGTC 960

CTAAGGCAGGATGTGGAGAGAGTTTGTAGTATAGGAACAGAGGAGAGACAGAGACTGTGCTACTGGTACTTCGATGTCTGGGGCGCAGGGACCAACCGCTCTCCTCAGGTAAGCTGG 1080

Fig. 4 Nucleotide sequence of the 5'-flanking region of J_{H1} DNA. Strategy for the sequencing is shown in Fig. 2Bc. The J_{H1} sequence and some of the 5'-flanking sequence have been published elsewhere^{12,13,43}. The D_{O52} and J_{H1} segments are underlined. Two blocks of conserved sequences.

GGTTTTTGTT and CACTGTG
CCAAAAACA GTGACAC

or their related sequences are boxed. Numbers indicate the lengths of the spacer separating the two conserved sequences in nucleotide bases.

ing is after Kabat *et al.*²⁴). The myeloma sequence of the 3'-flanking region of J_{H2} is identical to the embryonic sequence (data not shown). Thus, the right end of deletion was identified between the second and third nucleotides in the Asp codon (Fig. 3b). We then examined the sequence preceding the J_{H2} on the myeloma clone. As shown in Fig. 3b, a palindromic heptamer, CACAGTG, and a T-rich sequence, GGTTTTGAC, closely related to the nonamer GGTTTTGT, are present 14 and 26 base pairs, respectively, 5' to the recombination site. The spacer separating the heptamer and the nonamer is 12 base pairs long. These structural features of the DNA attached to the Q52 J_{H2} sequence fit the predictions made for the D segments. We could not find any V-like sequence on the 5' side of the recombination site, and this would have been easy to detect, as the sequence at the end of V DNA coding for residues 90-94 is highly conserved²¹.

Identification of a germ-line D segment

If the rearrangement on the Q52 clone represents the deletion between a D segment and a J_H segment, the germ-line D sequence must be on the 6.4-kilobase embryonic J_H fragment. As shown in Fig. 2, the left end of the deletion on the Q52 clone has been mapped 0.7 kilobases 5' to the J_{H1} DNA on the germ-line clone. We therefore extensively sequenced the 5'-flanking region of the J_H DNA cluster on the embryonic clone. Figure 4 shows the complete nucleotide sequence of the 1-kilobase region upstream from J_{H1} DNA. As predicted, about 700 base pairs 5' to the J_{H1} we identified the same sequence as found on the 5' side of Q52 J_{H2} DNA. In this germ-line sequence another set of the palindromic heptamer and the T-rich nonamer are invertedly repeated on the 3' side (Figs 4 and 3a). Again the spacer separating the heptamer and the nonamer is 12 base pairs. Accordingly, this germ-line sequence has the structure predicted for the germ-line D segment. In the D-like segment identified here, a 10-base pair DNA stretch is flanked by two palindromic heptamers.

CACAGTG CACCGTG
GTGTCAC GTGGCAC

and further flanked by A + T-rich nonamers.

GGTTTGGAC and ACAAAAACC
CCAAAAC TG TGTTTTGG

12 base pairs outside. This 10-base pair segment can code for three amino acid residues, Gln-Leu-Gly, Asn-Trp-Asp or Thr-Gly-Thr. Although no published mouse HV3 sequence cor-

responds to these hypothetical peptide sequences, we tentatively refer to this potential D segment as D_{O52}.

The tetranucleotides, AAGG, adjacent to the recombination site on the Q52 J_{H2} DNA cannot be explained by either the germ-line J_{H2} or the D_{Q52} sequence (Fig. 3), and at present the origin of this sequence is unknown. We looked for another germ-line D segment encoding the tetramer between the D_{Q52} and J_{H1} sequences and found none (Fig. 4). As shown in Fig. 3a, the coding sequence in the embryonic D_{Q52} segment is followed by pentanucleotides CACGG. If we assume one nucleotide deletion and one nucleotide substitution in this pentamer, the tetramer AAGG could be generated. These events might have occurred somatically on recombination (see the accompanying article¹⁷ for more discussion on this point).

In the hope of finding another D segment, we extensively sequenced the 1.3-kilobase segment 5' to D_{Q52} according to the strategy shown in Fig. 2Bc (sequence not shown). However, detection of another D or D-like sequence was unsuccessful: therefore, germ-line D segments do not seem to be as tightly clustered as are J_H DNAs.

Generality of the 12/23-base pair spacer rule

The present study extended the 12/23-base pair spacer rule^{12,13} to germ-line D segments. It thus seems that all recombinations leading to formation of a complete V gene are mediated by the same or similar enzymes. These enzymes are thought to contain two DNA-binding proteins, one for a DNA segment carrying the heptamer and nonamer separated by a short (12-base pair) spacer and the other for a DNA segment having essentially the same short sequences separated by a long (23-base pair) spacer^{12,13}. As the two pairs of the putative recognition sequences are in an inverted orientation relative to the corresponding coding DNA segments, they could form a transient intermediary duplex if it is stabilized by the recombinase^{9,13}.

D-J_H joining accompanies a deletion

It seems to be a general rule that sequence rearrangement involving immunoglobulin genes accompanies a deletion of the DNA sequence which separates a pair of joining DNA segments on a germ-line genome. Experimental evidence for this rule was first obtained for V-J joining of the λ_1 -type light-chain gene⁹, and was confirmed in a subsequent experiment²⁵ in the κ -type gene. There is also evidence that deletions accompany heavy-chain joining^{18,26}. Furthermore, a number of groups reported results which strongly suggest that the same rule applies to the 'switch recombination' of heavy-chain genes^{27-30,44}. All the evidence is based on hybridization analysis of mouse DNA using

an appropriate probe. The heteroduplex molecule, formed between the embryonic J_H clone and the Q52 J_H clone, provides the most direct evidence for the involvement of deletion in D- J_H joining. As the size of the cloned DNA fragments corresponds well to those of the fragments detected by Southern gel blotting analysis, the deletion cannot be attributed to cloning artefacts (see Fig. 1 for clone Q52 J_H).

D- J_H joining and allelic exclusion

Unlike all previous autosomal gene studies, only the gene on one of the two homologous chromosomes is active in a given cell for immunoglobulin chains³¹⁻³³. This 'allelic exclusion' is manifested in several ways for light chains. A common form is that one copy of the chromosome contains a rearranged, complete gene, and the other retains the germ-line configuration⁴. Alternatively, both alleles are rearranged, but one abortively in several different forms (refs 5, 18, 33-36, 45-48 and P. Kennedy and S.T., unpublished). At the heavy-chain locus, rearrangements on both copies of the chromosome seem to be very common (Fig. 1), and lack of rearrangement as observed in light chains does not seem to be a common form of allelic exclusion. The demonstration that the heavy-chain locus includes D DNA segments suggests a unique way in which this locus can be abortively rearranged. As shown here for myeloma QUPC52,

the unexpressed chromosome contains a D-J segment. Although allelic exclusion at the heavy-chain locus could be manifested by other forms of abortive rearrangement observed in the light-chain loci, the existence of D segments may provide additional forms. One might imagine that in addition to D-J joining V-D joining can occur.

The analysis of myeloma DNA indicates that rearrangement in the vicinity of J DNA segments is much more frequent in heavy-chain genes than in light-chain genes. This distinction also applies to normal B cells: little, if any J_H was observed in the germ-line configuration³⁷, whereas unrearranged J_K is observed in significant amounts³⁸. The reason for this difference is not clear. However, it is reasonable to assume that the frequency of joining of two separate DNA segments increases as the number of segments increases and/or as the distance between them decreases. If there are many D segments (see the accompanying article¹⁷) and if they are close to J segments as D_{Q52} is, D-J joining could account for more frequent abortive rearrangement in the heavy-chain locus.

We thank Lisa Sultzman, Lena Angman, Patricia Riegert and André Trauneker for technical assistance, Richard Maki and William Roeder for reading the manuscript, and Michael Potter for myelomas. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche & Co.

Received 7 January; accepted 11 March 1981.

- Hozumi, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3628-3632 (1976).
- Brack, C. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5652-5656 (1977).
- Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3171-3175 (1978).
- Brack, C., Hiram, M., Schuller, R. & Tonegawa, S. *Cell* **15**, 1-14 (1978).
- Lenhard-Schuller, R., Hohn, B., Brack, C., Hiram, M., & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4709-4713 (1978).
- Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15**, 1133-1144 (1978).
- Seidman, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3881-3885 (1978).
- Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280**, 370-375 (1979).
- Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
- Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
- Maki, R., Trauneker, A., Sakano, H., Roeder, W. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2138-2142 (1980).
- Early, P. *et al. Cell* **19**, 981-992 (1980).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
- Bernard, O. & Gough, N. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3630-3634 (1980).
- Honjo, T. *et al. Cold Spring Harb. Symp. quant. Biol.* (in the press).
- Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. *Nature* **283**, 35-40 (1980).
- Kurosawa, Y. *et al. Nature* **290**, 565-570 (1981).
- Cory, S., Adams, J. M. & Kemp, D. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4943-4947 (1980).
- Maki, R., Kearney, J., Paige, C. & Tonegawa, S. *Science* **209**, 1366-1369 (1980).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Yoshida, T. H., Imai, H. T. & Potter, M. *J. natn. Cancer Inst.* **41**, 1083-1089 (1968).
- Hengartner, H., Meo, T. & Müller, E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4494-4498 (1978).
- Tiemeier, D. C., Enquist, L. & Leder, P. *Nature* **263**, 526-528 (1976).
- Kabat, E. A., Wu, T. T. & Bilofsky, H. in *Sequences of Immunoglobulin Chains* (NIH, Bethesda, 1979).
- Seidman, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6022-6026 (1980).
- Cory, S. & Adams, J. M. *Cell* **19**, 37-51 (1980).
- Honjo, T. & Kataoka, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2140-2144 (1978).
- Yaoita, Y. & Honjo, T. *Nature* **286**, 850-853 (1980).
- Rabbitts, T. H. *et al. Nature* **283**, 351-356 (1980).
- Cory, S. *et al. Nature* **285**, 450 (1980).
- Pernis, B. *et al. J. exp. Med.* **122**, 853-875 (1965).
- Cebra, J. S., Colberg, J. E. & Dray, S. J. *J. exp. Med.* **123**, 547-558 (1966).
- Max, E. E. *et al. Cell* **21**, 793-799 (1980).
- Seidman, J. G. & Leder, P. *Nature* **276**, 790-795 (1978).
- Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937-1941 (1980).
- Seidman, J. G. & Leder, P. *Nature* **286**, 779-783 (1980).
- Nottenberg, C. & Weissman, I. L. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Joho, R. & Weissman, I. L. *Nature* **284**, 179-181 (1980).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
- Sakano, H. *et al. Nature* **277**, 627-633 (1979).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 498-502 (1980).
- Newell, N., Richards, J. E., Tucker, P. W. & Blattner, F. *Science* **209**, 1128-1132 (1980).
- Coleclough, C., Cooper, D. & Perry, R. P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1422-1426 (1980).
- Altenburger, W., Steinmetz, M. & Zachau, H. G. *Nature* **287**, 603-607 (1980).
- Choi, E., Kuehl, M. & Wall, R. *Nature* **286**, 776-779 (1980).
- Schnell, H. *et al. Nature* **286**, 170-173 (1980).
- Steinmetz, M., Zachau, H. G. & Mach, B. *Nucleic Acids Res.* **6**, 3213-3229 (1979).

Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes

Yoshikazu Kurosawa, Harald von Boehmer, Werner Haas, Hitoshi Sakano, André Trauneker & Susumu Tonegawa

Basel Institute for Immunology, 487 Grenzacherstrasse, Postfach, 4005 Basel 5, Switzerland

The finding that the diversity (D) and joining (J_H) but not the variable (V_H) DNA segments of mouse immunoglobulin heavy-chain genes are joined in the DNA of some cloned cytolytic T cells, led to identification and sequencing of three different D DNA segments. Two segments identified on the embryo DNA carry on both the 5' and 3' sides two sets of characteristic sequences separated by a 12-base pair spacer, which have been implicated as recognition signals for a recombinase. The third segment, identified in a form joined with a J_H DNA segment in a T cell, carries the recognition signal on the 5' side. These results support the 12/23-base pair model for somatic generation of immunoglobulin V genes, and rule out the possibility that the cytolytic T cells use assembled V_H , D and J_H sequences to encode their antigen receptors.

THE variable region of an immunoglobulin chain is encoded in multiple DNA segments scattered along a chromosome of a

germ-line genome¹⁻¹⁰. These DNA segments are assembled into a continuous stretch with concomitant deletion of the spacer

sequences during differentiation of B lymphocytes. The somatic rearrangement of DNA sequences is considered to have key roles in both somatic amplification of antibody diversity and determination of the specificity of lymphocyte clones. In mouse light chains of both λ and κ types, two DNA segments, V and J, are separate in the germ-line genome and are joined in the lymphocytes which express these DNA segments¹⁻⁵. The V and J segments code for about 95 residues at the NH₂-terminal and about 13 residues at the CO₂H-terminal ends of the V region, respectively^{1,3-5}. In contrast, DNA sequencing analysis of heavy-chain genes suggests that the V_H region is encoded by three DNA segments referred to as V_H, D and J_H, and somatic joining of the three DNA segments is necessary to generate a complete heavy-chain V gene⁶⁻⁸. The D DNA segment encodes the third hypervariable region, which determines in part the shape and size of the antigen-combining site^{11,12}.

The abovementioned sequence rearrangement of immunoglobulin genes should occur during the development of B lymphocytes before the appearance of the first detectable immunoglobulins. Does an analogous or similar rearrangement occur in the precursors of T lymphocytes, which share a common stem cell with B lymphocytes? This question is particularly important because of the possible use of the immunoglobulin V gene pool by T cells for coding their antigen-specific receptors¹³⁻¹⁶. One approach to this question is to analyse DNA of cloned T cells by the Southern gel blotting procedure using a J DNA sequence as the hybridization probe. A V-J or V-D-J joining will be monitored by a difference in the size of DNA fragments detected in embryo (or kidney) and T-cell DNA. We have therefore applied this technique to several cloned cytolytic T cells, and show here that sequence rearrangements frequently occur in T cells in the vicinity of the J_H sequences. Similar findings have been reported by two other groups with T-lymphoma lines¹⁷ and cytolytic T-cell lines¹⁸. However, our further analysis of two T-cell clones by recombinant DNA techniques demonstrates that the rearrangements observed in these T-cell clones are D-J joinings rather than V-D-J joinings. This led to the identification of two D DNA segments on the germ-line genome. One of the three D DNA segments described here was found to be identical to the D segment that was independently identified in our laboratory during the parallel analysis of an abortively rearranged sequence observed in the myeloma QUPC-52 (ref. 19).

T-cell DNA sequence is frequently rearranged in the vicinity of the J_H cluster

We have previously identified a cluster of four J_H DNA segments⁷ on a 6.4-kilobase *Eco*RI fragment cloned from BALB/c embryos (clone MEP203)⁶. A 1.1-kilobase *Sac*I-*Eco*RI fragment derived from the 3' end (relative to orientation of transcription) of this fragment (a J_H probe) would detect, when used as a hybridization probe on Southern blots of cellular DNA digested with *Eco*RI, sequence rearrangements that have occurred within the 6.4 kilobase *Eco*RI fragment. The results of various cytolytic T-cell clones and three thymomas are shown in Fig. 1a. The cytolytic T-cell clones, except B6.1, were confirmed for their functional activity before extraction of DNA. The embryo or kidney DNAs of five different mouse strains—BALB/c, C3H, C57BL/6, AKR and DBA—all gave the expected fragment of 6.4 kilobases previously identified. The DNA of three cytolytic T-cell clones—DAS 1, D.FL.2, D.FL.13—and two thymomas—BW 5147 and T5—each gave a single band indistinguishable from that of embryos or kidneys, suggesting that no gross sequence rearrangement has occurred in these cells within the 6.4-kilobase *Eco*RI fragment. [The faint bands observed above the 6.4-kilobase band in D.FL.2 and D.FL.13 DNA are due to incomplete digestion as demonstrated by separate experiments (data not shown).] In contrast, DNA of six cytolytic T cells—C.SP.2, C.FL.1, B6.1, D.FL.1, D.FL.12 and D.FL.16—gave one additional band of 5.5, 5.3 or 4.0

Table 1 Germ-line and myeloma D sequences

<u>Germ-line D</u>	
D _{SP2.1}	TC <u>TACTATGGTAAC</u>
D _{SP2.2}	TC <u>TACTATGATTACG</u> AC
D _{Q52}	C <u>AACTGGGAC</u>
<u>Myeloma D</u>	
S107	<u>TACTACGGTAGTA</u>
M603	<u>TACTACGGTAGTA</u> C
M167	G <u>GACTACGGTAATA</u> GCTACTTTG
MC101	AGG <u>GACTACGTTAGTA</u> GGTACGACCC
MPC11	GGGATT <u>TACTACAATAGTA</u> GCCC
M141	CGTCTCAATTTAT <u>TACTACGGTCGTA</u> GCGACAATACTTCACTT
M173-II	TGCCCTCTAC <u>TACTACAGTTACA</u> <u>GCGGGGGTT</u>
H76	<u>CCGGGGGTTCCCC</u>

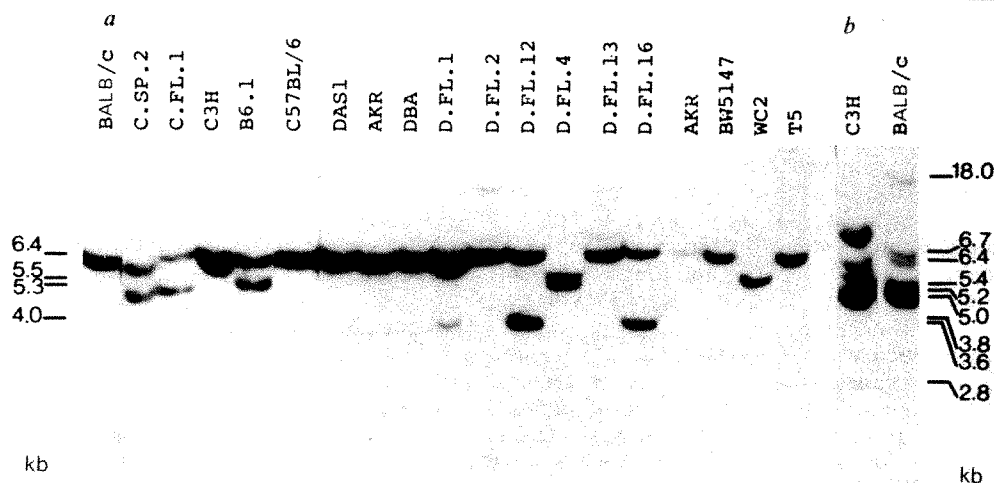
Three germ-line D sequences identified in the present study are compared with the D region sequences of the eight complete V_H genes present in the indicated myelomas. For D_{SP2.2} and D_{Q52} the sequences between the palindromic heptamers are shown. The D_{SP2.1} has been identified only in a form joined with the J_{H3} segment and therefore its exact 3' boundary is unknown. For S107 (ref. 8), M603 (ref. 8), M167 (ref. 8), MC101 (ref. 27) and M141 (ref. 7) the DNA sequences of both the complete (rearranged) and the embryonic (unrearranged) V_H genes are known. These sequences and the embryonic J_H sequences⁷ were used to deduce the 5' and the 3' boundaries of these D regions. For MPC11 (ref. 28), M173-II (H.S. and Y.K., unpublished) and H76 (ref. 9), only the DNA sequences of the complete V_H genes are known and therefore the 5' boundaries are ambiguous. The 3' boundaries were deduced as above. All sequences are aligned for the maximal homology. The D regions of S107 and M603 contain a 13-mer (underlined) which is a perfect palindrome around the central G. The homology present between each of the other D sequences and the 13-mer is also indicated by underlining. The homology between H76 and M173-II D sequences is indicated by double underlining.

kilobases. This indicates that some of the two or more copies of chromosome 12 [some T cells are polyploid (H.v.B. and W.H., unpublished)] in these T cells have undergone sequence rearrangement involving the 6.4-kilobase embryonic fragment. In one cytolytic T-cell clone, D.FL. 4, and one thymoma, WC2, no 6.4-kilobase embryonic fragment was observed. The presence of two non-embryonic bands, of 5.5 and 5.3 kilobases, in D.FL.4 suggests that sequence rearrangement has occurred on both (assuming diploidy) chromosomal copies in different ways. In contrast, WC2 seems to contain only one copy of the 6.4-kilobase sequence and sequence rearrangement has occurred in this fragment. Thus, we have observed sequence rearrangement in 8 out of 13 cytolytic T-cell or thymoma clones. In the majority, rearrangement occurred only in some but not all of the copies of chromosome 12. The sizes of the rearranged fragments seem to fall into three categories—5.5, 5.3 and 4.0 kilobases.

Two DNA clones from two cytolytic T cells

We chose B6.1 and C.SP2 for further studies and cloned the rearranged DNA fragments of 5.5 and 5.3 kilobases, respectively, using a λ WES vector²⁰; we refer to these fragments as B6B1 (short-term B6) and SP2B1 (short-term SP2), respectively. Electron microscopic examination of the heteroduplex molecules formed between either of these T-cell DNA clones and the 6.4-kilobase embryonic *Eco*RI fragment (clone ME183-8) allowed mapping of the recombination sites. The B6 and embryo DNA fragments gave a single-stranded DNA loop flanked by double-stranded DNA arms (Fig. 2a). In contrast, the SP2 and embryo DNA fragments gave Y-shaped heteroduplexes. Measurement of the various parts of the heteroduplex molecules indicated that the B6 DNA fragment has undergone a 1.2-kilobase deletion in a region immediately adjacent (5' side) to the J_{H2} sequence. In contrast, the entire sequence 5' to J_{H3} is

Fig. 1 *a*, Hybridization of T-cell and thymoma DNA with the J_H probe. High molecular weight DNA was prepared from various cloned cytolytic T cells and thymomas by the procedure described previously²⁹, digested to completion with *EcoRI*, and fractionated on 0.8% agarose gel by electrophoresis. The DNA was denatured and transferred *in situ* to nitrocellulose filters (Schleicher and Schuell, BA85) essentially as described previously³⁰. The filters were incubated for hybridization with nick-translated J_H probe (1.1-kilobase *SacI-EcoRI* fragment indicated in Fig. 3A) according to the published procedures². Characteristics of the various cytolytic cells have been described previously^{31,32}. C.SP.2 and C.FL.1 are derived from C3H, B6.1 from C57BL/6, DAS 1 from AKR/DBA F1, all D.FL clones from DBA, and thymomas were all derived from AKR. For comparison, DNA from BALB/c embryos and kidneys of C3H, C57BL/6, AKR and DBA were also analysed. These DNA are indicated by strain names. Numbers and arrows to the left indicate sizes of DNA fragments in kilobases (kb). *b*, Hybridization of BALB/c embryo and C3H kidney DNA with SP2D-J probe. A nitrocellulose filter, prepared as described for *a*, was incubated with the nick-translated SP2D-J probe, 0.8-kilobase *BglII-HindIII* fragment indicated in Fig. 3C. The sizes of the BALB/c DNA fragments detected are shown in kilobases. The 3.8-, 3.6- and 2.8-kilobase bands could be seen in the original autoradiogram, but are too faint to be reproduced.



replaced with a sequence of unknown origin in the SP2 DNA fragment. The points of sequence divergence between the T and embryo cell DNA fragments were further determined by comparison of restriction enzyme cleavage sites (Fig. 3A-C). The results confirm the conclusion drawn by heteroduplex mapping.

The SP2 fragment contains a joined D-J DNA segment

We determined the DNA sequence of the 0.8-kilobase *BglII-HindIII* fragment (Fig. 3Ca) of the SP2 fragment containing the site of recombination according to the strategy shown in Fig. 3Cb. This sequence was aligned with the previously determined sequence of the embryonic fragment in the J_H -coding region (Fig. 4). As suggested by the heteroduplex molecules (Fig. 2b) and the restriction enzyme maps (Fig. 3A, C), the sequences of the two DNA fragments match well starting with the first base of the Phe codon TTT located in the 5' edge of the J_H DNA segment. On the SP2 fragment the Phe codon is preceded by an Asn codon, AAC, which is in turn preceded by a nonamer, TACTATGGT, that codes for a trimeric peptide, Tyr-Tyr-Gly, when read in phase with the J_H . This nonameric nucleotide sequence or its close variants appear in the D regions of seven out of eight myeloma V_H genes sequenced to date (Table 1). Furthermore, the trimeric peptide, Tyr-Tyr-Gly, encoded by the nonamer has been found in the D regions of a majority of myeloma V regions for which amino acid sequence data are available¹¹. These data strongly suggest that a D DNA segment is joined with the J_H sequence on the SP2 fragment.

Further evidence for this notion comes from the feature of the nucleotide sequence located immediately 5' to the putative D DNA segment on the SP2 fragment. Based on the DNA sequencing analysis of the various germ-line DNA segments, it has been predicted that a nonamer,

GGTTTTTGT
CCTAAAACA

and a palindromic heptamer,

CACAGTG
GTGTCAC

separated by a 12-base pair spacer or closely related variants of these sequences, precedes a germ-line D DNA segment^{7,8}. As shown in Fig. 4, the putative D segment joined with the J_H on the SP2 fragment is preceded by two sequences,

GATTTTTGT and TACTGTG
CTAAAACA ATGACAC

separated by a 12-base pair spacer. These results partially fulfill

the predictions of the 12/23-base pair spacer model^{7,8}, and in turn support the conclusion that the sequence joined with the J_H on the SP2 fragment is indeed a D DNA segment. We refer to this D DNA segment as $D_{SP2.1}$.

No V_H -D- J_H joining on the SP2 fragment

Is a germ-line V_H DNA segment also joined on the SP2 fragment? A visual inspection of the SP2 DNA sequence in the region preceding the $D_{SP2.1}$ segment indicates that no germ-line V_H or V_H -like sequence is attached to the D DNA segment. First, there are one, two or four translation termination codon(s) in the three different reading frames of the 300-base pair (an approximate size of a germ-line V_H segment) region 5' to the $D_{SP2.1}$ segment. Thus, this region cannot constitute part of a

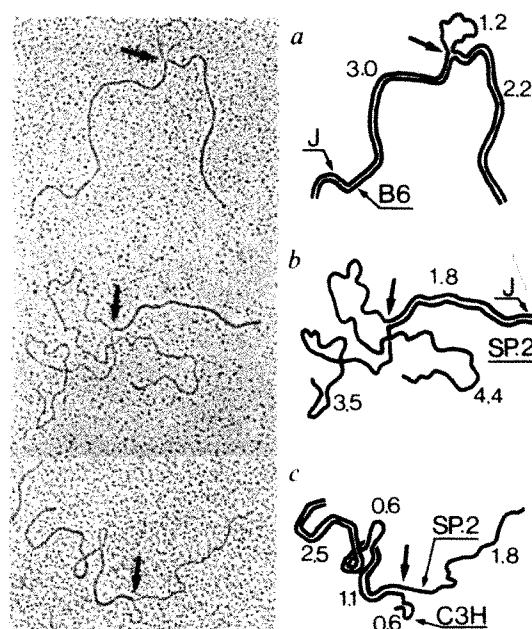


Fig. 2 Heteroduplex molecules formed between cloned DNA fragments. Cloned DNA was digested with *EcoRI* and examined under an electron microscope as described previously². Thick arrows indicate points at which heteroduplexes branch out. The numbers indicate lengths of various parts of the molecules in kilobases. In this experiment we used clone ME184-8 instead of MEP203^{6,7} as a source of the 6.4-kilobase embryonic *EcoRI* fragment. Clone ME184-8 was prepared by Richard Maki and contains, in a phage vector, λ_{WES} , the 6.4-kilobase *EcoRI* fragment carrying the four J_H DNA segments. *a*, 6.4-kilobase ME184-8 (J in short-form) versus 5.3-kilobase B6B1 (B6); *b*, 6.4-kilobase ME184-8 (J) versus 5.5-kilobase SP2B1 (SP2); *c*, 5.5-kilobase SP2B1 (SP2) versus 5.0-kilobase C3H-2 (C3H).

germ-line V_H DNA segment. Second, some amino acid residues are highly conserved among the V_H regions. In particular, Cys residues are invariably present in all species studied to date at positions 22 and 92 and are the only Cys residues in the V_H region¹⁰. These two Cys residues play an important part in folding the protein domains through a disulphide linkage^{21,22}. In addition, the Trp at position 36 is invariant across species¹¹. The 300-base pair SP2 DNA sequence preceding the $D_{SP2.1}$ segment contains three Cys and one Trp codons in the frame defined by the J_{H3} peptide. Eight Cys and four Trp codons, and four Cys and one Trp codons, respectively, exist in the other two reading frames. Thus, in all cases the total number of Cys codons is incompatible with the notion that the DNA sequence before the D DNA segment comprises a V_H DNA segment. Furthermore, the locations of Cys and Trp codons within this DNA sequence bear no resemblance to those in a V_H DNA segment. We conclude that no V_H or V_H -like DNA segment is joined with the $D-J_H$ DNA on the SP2 fragment.

Identification of a $D_{SP2.1}$ -like D DNA segment in a germ-line genome

Identification of a $D-J_H$ joining in the DNA of a cytolytic T-cell clone allowed the identification, cloning and characterization of germ-line D DNA segments. We dissected out the 0.8-kilobase *Bgl*II-*Hind*III fragment (Fig. 3C) from the SP2 clone and used it as a hybridization probe (SP2DJ probe) for analysis of *Eco*RI-digested total DNAs of both BALB/c and C3H embryos (Fig. 1b). As this probe contains the J_{H3} and its 3'-flanking sequence, it detects the 6.4-kilobase embryonic *Eco*RI fragment carrying

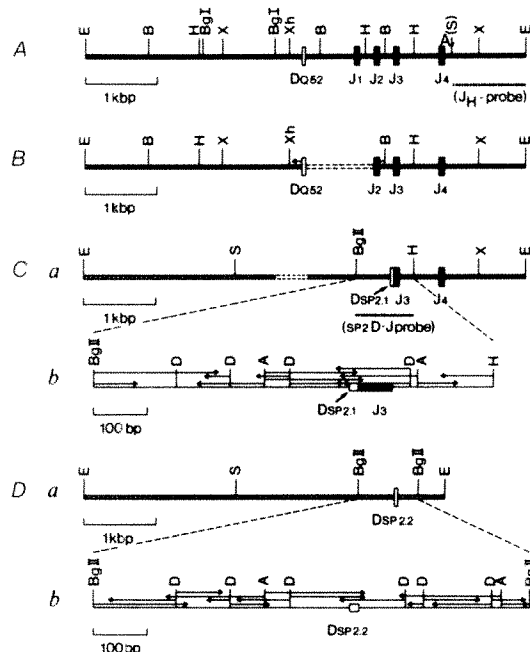


Fig. 3 Restriction enzyme cleavage maps of the *Eco*RI inserts of various DNA clones. **A**, The 6.4-kilobase *Eco*RI fragment of MEP203^{6,7}; **B**, B6B1; **C**, SP2B1; **D**, C3H-2. Clone MEP203, which was described previously^{6,7}, is an embryonic DNA clone and contains three *Eco*RI fragments, one of which is the 6.4-kilobase fragment carrying the four J_H segments. Cleavage sites were determined by single and double digestions of the cloned DNA by various restriction enzymes. Homologous regions are indicated by bars of two types of shading. The broken lines in **B** indicate the 0.6-kilobase region that appeared as a deletion loop in the heteroduplex molecules formed between the SP2B1 and C3H-2 fragments. The filled boxes correspond to the four J_H DNA segments previously identified on the MEP203 6.4-kilobase fragment^{6,7}. The open boxes are D segments identified by DNA sequencing. Sequencing strategy and finer restriction maps of the 0.8-kilobase *Bgl*II-*Hind*III fragment from SP2B1 and the 0.85-kilobase *Bgl*II-*Hind*III fragment from C3H-2 are shown in **Cb** and **Dd**, respectively. Horizontal arrows indicate the directions and extent of sequencing. The sequence across the 1.2-kilobase deletion of the B6B1 fragment was determined using the *Xba*I-*Bam*HI fragment (**B**). The J_H probe used in Fig. 1a and the SP2DJ probe used in Fig. 1b are indicated in **A** and **Ca**, respectively. Abbreviations used are: E, *Eco*RI; B, *Bam*HI; H, *Hind*III; BgI, *Bgl*II; X, *Xba*I; Xh, *Xho*I; S, *Sac*I; BgII, *Bgl*II; D, *Dde*I; A, *Acl*I; kbp, kilobase pairs; bp, base pairs.

the four germ-line J_H segments. In addition, the probe detected one (5.0 kilobase), four (18, 6.7, 5.4 and 5.2 kilobase) and three (3.8, 3.6 and 2.8 kilobase) bands of high, medium and low intensity, respectively, in BALB/c DNA. A similar, but distinctive hybridization pattern was obtained in C3H DNA. These results may indicate that a number of $D_{SP2.1}$ or $D_{SP2.1}$ -like DNA segments exist in the germ-line genomes of mice, some of which are polymorphic.

We have cloned most of these DNA segments in phage λ vectors and determined their structure and organization in the germ-line genome (Y.K., unpublished). Some characteristics of one of the most prominent fragments, that of 5.0 kilobases (clone C3H-2, C3H in short-term) isolated from C3H embryos, are described here. The heteroduplex formed between clone C3H-2 and the SP2 fragment (Fig. 2c) and the restriction enzyme maps (Fig. 3C, D) indicate that the two DNA fragments are extensively homologous in the 3.6-kilobase region which starts with the 5' *Eco*RI end and ends somewhere near the position of $D_{SP2.1}$ on the SP2 fragment. However, that the homology is imperfect is clearly shown by the presence of a 0.6-kilobase insert (or deletion) in the C3H (or SP2) clone. Nevertheless, the above result suggests that the clone C3H may contain a $D_{SP2.1}$ -like DNA segment at the 3' edge of the homology. We therefore prepared a finer restriction map around this region and determined the entire DNA sequence of the 0.85-kilobase *Bgl*II-*Bgl*II fragment, according to the strategy shown in Fig. 3Db.

As predicted, the sequence homology between the SP2 and C3H fragments is extensive, starting with the 5'-*Bgl*II site and ending with the triplet AAC, the last codon of $D_{SP2.1}$ on the SP2 fragment (Fig. 4). There are 28 single-base substitutions or deletions (or insertions) in the 485-base pair homology region, two of which are within the D DNA segments. The D DNA segment on the embryonic clone C3H is flanked by two sets of the putative recognition sequences with 12-base pair spacers. This confirms the universality of the 12/23-base pair spacer model^{7,8}. Although the C3H-2 D DNA segment and $D_{SP2.1}$ are highly homologous, the imperfectness of the homology in the 5'-flanking regions as well as in the coding segments themselves clearly indicate that they are two different copies of the D segments that occupy independent positions in the germ-line genome. We therefore refer to the C3H-2 D DNA segment as $D_{SP2.2}$.

The B6 fragment contains another joined D-J sequence

Figure 5 shows the sequence across the 1.2-kilobase deletion of the B6 fragment and sequences of the embryonic fragment that cover the 5' and 3' boundaries of the deletion. The sequence homology between the B6 and the embryonic fragments ends with the C-G base pair immediately to the left of the line Ld and resumes with a T-A base pair immediately to the right of line Rd. The latter base pair corresponds to the second letter of the Phe codon TTT located near the 5' terminus of the previously identified J_{H2} sequence^{7,8}. The single base pair C-G present on the B6 fragment and flanked by the two lines, Ld and Rd, is unaccounted for by the embryonic sequence (see below). Ld is preceded by a putative enzyme recognition sequence,

GGTTTTGAC (12-base pair spacer) CACAGTG
CCAAAAC TGTCAC

both on the B6 and MEP203 fragments. On the latter fragment Ld is followed by nine base pairs, which are in turn followed by another putative signal sequence,

CACGGTG (12-base pair spacer) ACAAAAACC
GTGCCAC TGTTTTTGG

Thus, the 9-base pair sequence on the MEP203 fragment fulfills the requirement imposed on a D DNA segment by the 12/23-base pair spacer rule^{7,8}. We therefore conclude that the B6 deletion is due to the joining of the J_{H2} sequence with a D DNA

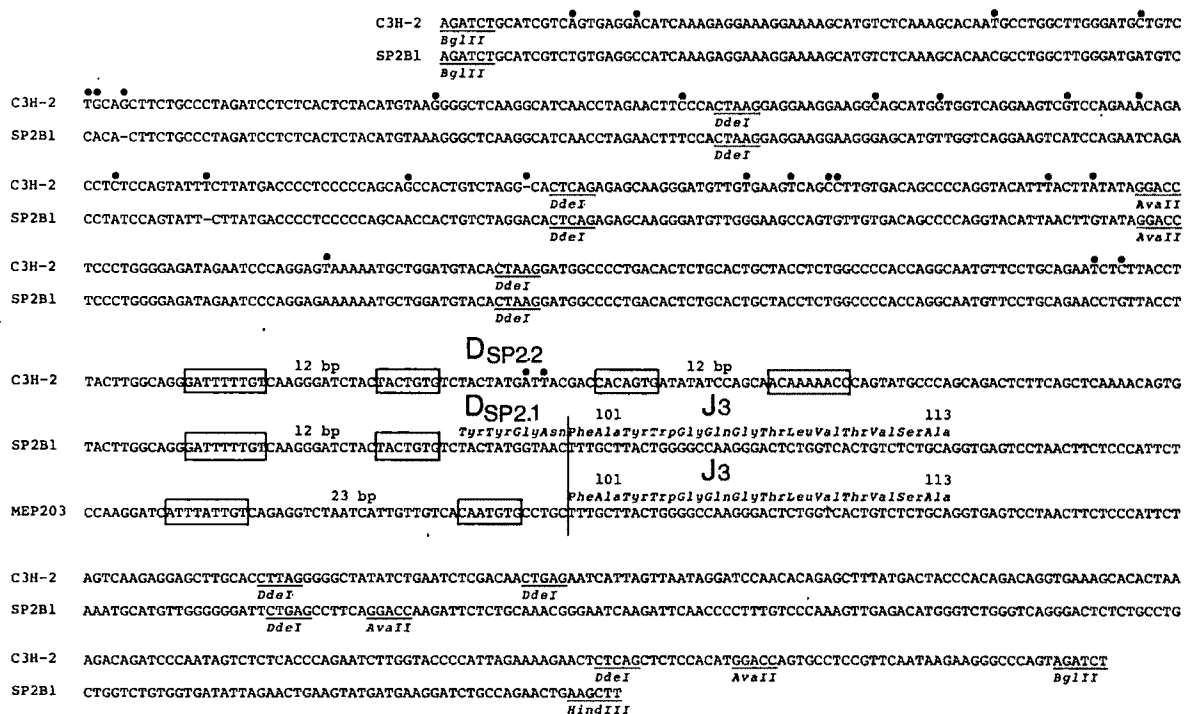


Fig. 4 Nucleotide sequences of three DNA clones: C3H-2, SP2B1 and MEP203 (refs 6, 7). All sequences were determined by the method of Maxam and Gilbert³³. The sequence of the 0.85-kilobase *Bgl*III–*Bgl*III fragment of clone C3H-2 is aligned for the maximal homology with the sequence of the 0.8-kilobase *Bgl*III–*Hind*III fragment of clone SP2B1. Also aligned with the SP2B1 sequence is the sequence of the *J*_{H3}-containing region of clone MEP203, which was determined previously in our laboratory⁷. The SP2B1 fragment contains a D segment (*D*_{SP2.1}) fused with the *J*_{H3} segment. Amino acid residues encoded by the fused *D*_{SP2.1} and *J*_{H3} on the SP2B1 fragment and those by the unfused *J*_{H3} on the MEP203, numbered according to Kabat *et al.*¹¹, are indicated in italics. The SP2B1 sequence is extensively homologous to the C3H-2 sequence from the 5' *Bgl*III site to the *D*_{SP2.1}–*J*_{H2} boundary, where the homology switches to the MEP203 sequence. The dots and dashes indicate base substitutions and deletions within the homologous regions. The putative recognition sequences for recombinations are boxed. C3H-2 contains a D segment, *D*_{SP2.2}.

segment located 1.2 kilobases upstream. This D DNA segment has been independently identified during analysis of a rearranged fragment isolated from myeloma QUPC52 and is therefore referred to as *D*_{Q52} (see the accompanying article¹⁹).

As mentioned above, one base pair, C·G, is inserted on the B6 fragment at the site of recombination. Based on repeated sequencing of the relevant region, we believe that error is unlikely. This suggested the possible involvement of an intermediate recombination, the extra base pair arising from an independent D DNA segment. However, as described in the accompanying article, no additional D sequence has been identified between *D*_{Q52} and *J*_{H2}. Thus, if one is to pursue this possibility, it is necessary to invoke a scheme that allows joinings of multiple D DNA segments, the order of which in the assembled gene is independent of that in the germ-line genome. Although such a scheme is possible, one alternative is that the act of recombination introduced the extra base pair at the site of joining.

Structure and diversity of D sequences

In the digest of BALB/c mouse embryo DNA (Fig. 1b) the SP2D probe detected at least eight *Eco*RI fragments, each of which is likely to contain at least one copy of a *D*_{SP2.1}-like DNA segment. For the 5.0-kilobase fragment, this was demonstrated directly by DNA sequencing, which led to the identification of *D*_{SP2.2}. Our recent result indicates that the 5.0-kilobase band contains three *Eco*RI fragments each of which hybridizes with the SP2DJ probe (Y.K., unpublished). Thus, assuming that positive hybridization with the probe signifies the presence of at least one D segment, the minimum number of *D*_{SP2.1}-like segments is 10 (8 bands, of which one is triple).

If *D*_{SP2.1}-like DNA segments are abundant, as the Southern gel blot suggests, this is in agreement with the frequent appearance of the *D*_{SP2.1}-like sequences in the D regions of the

assembled V genes. The D-coding regions of seven out of eight assembled *V*_H genes sequenced to date contain a palindromic 13-mer, TACTACGGTAGTA (ref. 8), or its variants (Table 1). The *D*_{SP2.1} and *D*_{SP2.2} also contain sequences similar to the 13-mer (Fig. 4, Table 1). In addition, the D regions of most mouse *V*_H regions for which amino acid sequences are available contain a tripeptide Tyr-Tyr-Gly that is encoded by the *D*_{SP2.1} and *D*_{SP2.2} segments¹¹. Even the *D*_{Q52} segment which did not cross-hybridize with the *D*_{SP2.1}-like segments (Y.K., unpublished) shows significant homology, not only in the coding region (Table 1), but also in the flanking regions (Figs 4, 5). Thus, many, if not all D segments are related to the *D*_{SP2} segments, both evolutionarily and structurally.

Many D regions have extensions on one or both side(s) of the 13-mer (Table 1). The only striking characteristic of the sequences within these extensions is that the 3' extension of the M173-II region shares a heptamer, CGGGGGT, with the H76 D region, the only D region apparently having no homology with the core 13-mer (Table 1). Each of these D regions may be encoded by its own D segment. Alternatively, some D regions, particularly those with relatively long extensions, may be encoded in two or more discrete D segments which are joined in the complete V genes. Two observations argue for the presence of D–D joinings. First, one and four base pairs at the presumed *D*_{Q52}–*J*_{H2} joining sites of the B6B1 and Q52 clones were in fact encoded neither by the *D*_{Q52} nor *J*_{H2} DNA segment (ref. 19 and Fig. 5). Second, we have been unable to detect a *D*_{M141} band in Southern gel blots of embryo DNA when this sequence of 44 base pairs (Table 1) was dissected out from the complete *V*_{M141} gene and used as the hybridization probe. In the same hybridization conditions the *J*_{H4} band was easily detected using the corresponding probe of nearly the same length (H.S., unpublished). If D–D joinings indeed occur, we predict that some D segments carry putative recognition sequences with 23-base pair spacers on at least one side.

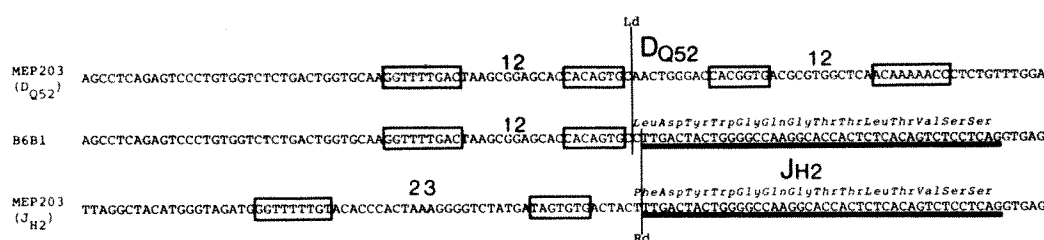
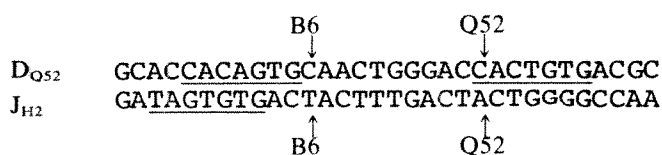


Fig. 5 Nucleotide sequences of the recombination regions of clone MEP203 and clone B6B1. The sequence across the 1.2-kilobase deletion of the B6B1 clone is compared with the two sequences of the embryonic fragment, MEP203, that cover the 5' (top)¹⁹ and the 3' (bottom)⁷ boundaries of the deleted segment. The two vertical lines, Ld and Rd, designate the left (5' side) and the right (3' side) boundaries of the deletion as determined by DNA sequencing, respectively. The other markings are as in Fig. 4.

Somatic diversification of heavy chains by V_H-D-J_H joinings

Amino acid sequence data¹² strongly suggest that the three types of germ-line V gene segments, V_H, D and J_H, combine in various, albeit not all possible, combinations to form complete V genes that are active in myelomas and presumably in B lymphocytes. The presence of the common 'recognition' sequences on the joining ends of all three gene segments supports this 'random' assortment. Furthermore, in both κ light- and heavy-chain genes it has been established that a given J DNA segment can assort with different V DNA segments^{4,5,23}.

Another somatic mechanism that might be used to amplify the coding capacity of immunoglobulin genes has been previously suggested by us⁴ and by Max *et al.*⁵, based on a proposed flexibility in the joining enzyme(s) with respect to the exact recombination sites on a given pair of joining gene segments. Although several previous observations in the κ -gene system support this hypothesis^{4,5,7,24,25}, none of the previous studies provides direct evidence for the hypothesis, because no two joinings examined took place between the same pair of gene segments. The present study, when combined with that described in the accompanying paper¹⁹, presents such a case. Thus, the D-J joinings observed in myeloma QUPC52 and T-cell B6 both used the D_{Q52} and J_{H2}, and yet the exact recombination sites differ both on the D and J sides in the two joinings as shown below,



where the arrows indicate the recombination sites and the underlined sequences are the palindromic heptamers. Provided that these recombinants represent the intermediates of successful V-D-J joinings, we can conclude that modulation of exact recombination sites is another source of antibody diversity that is generated somatically.

Received 7 January; accepted 11 March 1981.

1. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3171-3175 (1978).
2. Brack, C., Hiram, M., Schuller, R. & Tonegawa, S. *Cell* **15**, 1-14 (1978).
3. Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15**, 1133-1144 (1978).
4. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
5. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
6. Maki, R., Trautwein, A., Sakano, H., Roeder, W. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2138-2142 (1980).
7. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
8. Early, P. *et al. Cell* **19**, 981-992 (1980).
9. Bernard, O. & Gough, N. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3630-3634 (1980).
10. Rabbitts, T. H., Matthysens, G. & Hamlyn, P. H. *Nature* **284**, 238-243 (1980).
11. Kabat, E. A., Wu, T. T. & Bilofsky, H. in *Sequences of Immunoglobulin Chains* (NIH, Bethesda, 1979).
12. Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. *Nature* **283**, 35-40 (1980).
13. Binz, H. & Wigzell, H. *Contemporary Topics Immun.* **7**, 113-177 (1977).
14. Rajewsky, K. & Eichmann, K. *Contemporary Topics Immun.* **7**, 69-112 (1977).

T-cell antigen receptors are not encoded in assembled V_H, D and J_H segments

Our gel-blotting experiments using the J_H hybridization probe demonstrated that the DNA sequences are often, but perhaps more importantly not always, rearranged in the vicinity of the J_H cluster in various clones of cytolytic T cells and in thymomas. Two other groups reported similar findings and implicated them in the somatic formation of the complete structural genes coding for the antigen receptors of T cells^{17,18}. The present results clearly rule out such a simple interpretation of the gel-blotting experiments. The recombinations in both SP2 and B6 cells involve D and J sequences, but no V_H sequence. In fact, neither recombinant can code for a V-like polypeptide chain in the region located 5' to the recombination sites. We also emphasize that the recombination does not always occur among T-cell clones; we observed no indication of rearrangement around the J_H cluster in four well characterized clones of functional cytolytic T cells. We thus conclude that if the V_H DNA pool is shared by T and B cells to encode the antigen receptors, the four J_H segments identified to date are not used for rearrangement, at least in cytolytic T cells. It may be that T-cell receptors are encoded in an entirely different set of genes, or alternatively T and B cells share the same or an overlapping set of germ-line V_H genes, and another set of J-like and C-like DNA segments exists elsewhere in the genome, with which a V DNA segment joins specifically in T cells. Why, then, are there D-J_H joinings in T cells? One possibility is that they occurred in the common precursor cell which gave rise to B and T cells and have no active role in T-cell function. As discussed in the accompanying article¹⁹, D-J_H joinings might be common on the unexpressed copy of chromosome 12 in myelomas, and one recent study²⁶ suggests that this might also be the case in natural B lymphocytes. Alternatively, the hypothetical T-cell recombinase may mediate these joinings due to lack of absolute specificity for the gene segments encoding the T-cell receptors.

We thank Lena Angman, Lisa Sultzman, Patricia Riegert and Ghulamreza Dastoornikoo for technical assistance, Michael Julius and Bonnie Blomberg for discussion, Michael Potter for myelomas, Richard Maki for DNA clone ME184-8, Marcus Nabholz for T-cell clone DAS 1 and A. Coutinho for thymomas. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche & Co., Ltd, Basel, Switzerland.

15. Tada, T. & Okumura, K. *Adv. Immun.* **28**, 1-87 (1979).
16. Julius, M. H., Cosenza, H. & Augustin, A. A. *Nature* **267**, 437-439 (1977).
17. Cory, S., Adams, J. M. & Kemp, D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4943-4947 (1980).
18. Forster, A., Hobart, M., Hengartner, H. & Rabbitts, T. H. *Nature* **286**, 897-899 (1980).
19. Sakano, H., Kurosawa, Y., Weigert, M. & Tonegawa, S. *Nature* **290**, 562-565 (1981).
20. Leder, P., Tiemeier, D. & Enquist, L. *Science* **196**, 175-177 (1977).
21. Edelman, G. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **63**, 78-85 (1969).
22. Poljak, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **70**, 3305-3310 (1973).
23. Weigert, M., Gattmann, L., Loh, E., Schilling, J. & Hood, L. *Nature* **276**, 785-790 (1978).
24. Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280**, 370-375 (1979).
25. Max, E. E., Seidman, J. G., Miller, H. & Leder, P. *Cell* **21**, 793-799 (1980).
26. Nottenberg, C. & Weissman, I. *Proc. natn. Acad. Sci. U.S.A.* **77**, in the press (1980).
27. Honjo, T. *et al. Cold Spring Harbor Symp. quant. Biol.* (in the press).
28. Zakut, R., Cohen, J. & Givol, D. *Nucleic Acids Res.* **8**, 3591-3601 (1980).
29. Hozumi, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3628-3632 (1976).
30. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
31. Haas, W., Mathur-Rochat, J., Pohlitz, H., Nabholz, M. & von Boehmer, H. *Eur. J. Immun.* **10**, 828-834 (1980).
32. von Boehmer, H. *et al. Eur. J. Immun.* **9**, 592-597 (1979).
33. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1980).

LETTERS

A lower limit for the birth rate of pulsars

Ramesh Narayan & M. Vivekanand

Raman Research Institute, Bangalore-560080, India

Using experimental data on observed pulsars, we have estimated a lower limit for the birth rate of pulsars in our Galaxy to be $B_{1b} = 0.0111 (\pm 0.0050) K$ pulsars yr^{-1} , where K , the 'beam factor', allows for the possibility that only a fraction of all pulsars is beamed towards the Earth. If we take K to be 5, we obtain a mean birth rate of 1 pulsar per 18 yr while we can place a firmer lower bound with 95% confidence of 1 per 50 yr. Two important new features of the present approach are: (1) we do not assume any model for pulsars; (2) we have allowed for selection effects (in a model independent way). Apart from obtaining a more reliable number for the birth rate of pulsars, we conclude that the quantity $P/(2\dot{P})$, where P is the period of the pulsar and $\dot{P}$ is its time derivative, is probably a good measure of age in 'young' pulsars up to half a million years old, and that selection effects which have been neglected in all earlier calculations are significant. The present calculation reduces the discrepancy between pulsar and supernova birth rates.

We assume that $\dot{P}$ is always positive and that the pulsar distribution in the Galaxy is in a steady state. Let $\rho_i(P, \dot{P}, L) dP d\dot{P} dL$ be the density distribution of pulsars in the Galaxy, where L is the intrinsic pulsar radio luminosity. The distribution $\rho_o(P, \dot{P}, L)$ of pulsars observable from the Earth is then given by

$$\rho_i(P, \dot{P}, L) = KS(L)\rho_o(P, \dot{P}, L) \quad (1)$$

where the scale factor $S(L)$ is the ratio of the total volume of the Galaxy to the volume covered in pulsar searches, both weighted by the known radial density distribution of pulsars in the Galaxy. K is assumed to be independent of P , $\dot{P}$ and L and to have a value¹ 5. Now, the flux or current of pulsars at any period P moving from lower to higher periods is given by

$$\begin{aligned} f(P) &= \int_0^\infty \int_0^\infty \rho_i(P, \dot{P}, L) \dot{P} d\dot{P} dL \\ &= K \int_0^\infty \int_0^\infty S(L) \rho_o(P, \dot{P}, L) \dot{P} d\dot{P} dL \end{aligned} \quad (2)$$

Clearly, $f(P)$ is equal to the number of pulsars being born per unit time in the period $0-P$, minus those dying in the same range. Therefore, $f(P)$ is rigorously less than or equal to the total birth rate B . For better statistics, we consider the following average of $f(P)$

$$F(P_{\max}) = \frac{1}{P_{\max}} \int_0^{P_{\max}} f(P) dP \leq B \quad (3)$$

where the rigorous inequality is obviously true for all values of $P_{\max}$.

We have estimated $F(P_{\max})$ by means of equation (4) using experimental data available for 210 pulsars detected by the Jodrell Bank², Arecibo³ and Second Molonglo⁴ surveys above their quoted minimum flux sensitivities:

$$F_{\text{mean}}(P_{\max}) = \frac{K}{P_{\max}} \sum_i S(L_i) \dot{P}_i; \quad 0 < P_i \leq P_{\max} \quad (4)$$

The quantities $S(L_i)$ were computed using published experimental details regarding sky coverage, receiver temperature, and so on of the above three surveys. As a check on our values of

$S(L_i)$, we have calculated the total number ($\pm$ standard deviation) of pulsars in the Galaxy to be

$$N = K \sum_i S(L_i) = 6.05 (\pm 1.88) \times 10^5 \text{ pulsars} \quad (5)$$

From a careful statistical analysis we have also obtained a lower limit N_{951b} with 95% confidence to be 3.19×10^5 pulsars. These numbers are in good agreement with the standard quoted value^{1,5} of 5×10^5 pulsars. Using equation (4), we have calculated $F_{\text{mean}}(P_{\max})$ and also $F_{951b}(P_{\max})$ at various values of $P_{\max}$ (the solid and dashed lines in Fig. 1). By equation (3), any value of $P_{\max}$ can be selected to set a lower bound on B . However, F_{mean} is sensitive to statistical fluctuations arising from individual pulsars whereas F_{951b} is relatively unaffected. We have therefore selected the $P_{\max}$ at which F_{951b} is maximum. Thus

$$\begin{aligned} B_{1b} &= F_{\text{mean}}(1.3 \text{ s}) \\ &= 0.0555 (\pm 0.0248) \text{ pulsars per yr per Galaxy} \\ &= 1 \text{ pulsar in } 18_{-6}^{+15} \text{ yr} \end{aligned} \quad (6)$$

$$\begin{aligned} B_{951b} &= 0.0200 \text{ pulsars per yr per Galaxy} \\ &= 1 \text{ pulsar per 50 yr} \end{aligned}$$

The present calculation has two distinct advantages over earlier approaches.

All earlier attempts to estimate B assume^{1,2} that the age τ of a pulsar is given by the quantity $P/(2\dot{P})$. This relation implies that pulsars slow down because of magnetic dipole radiation⁶ and that the magnetic field does not decay. Our calculation has no such model dependent assumptions. However, we have also developed a variant of the τ calculation⁷ using our concept of pulsar current and individual scales, $S(L_i)$. This gives a B_{1b} of 1 pulsar per 27_{-9}^{+29} yr and a B_{951b} of 1 pulsar per 84 yr. From the approximate agreement of these numbers with our rigorous results in equation (6) we conclude that $P/(2\dot{P})$ is a reasonable measure of pulsar age up to about half a million years, thus setting a lower limit on the decay time of the magnetic field.

This is the first calculation to allow for any selection effects which might arise from pulsar luminosities being correlated with P or $\dot{P}$. Such correlations have been neglected in earlier calculations in spite of there being some evidence⁸ for them. To assess the significance of selection effects, we have computed B_{1b} using equation (4) with a single average scale factor $\langle S(L) \rangle$ for all pulsars⁷, thus in essence neglecting correlations between L and P or $\dot{P}$. We then obtain a birth rate of 1 pulsar per

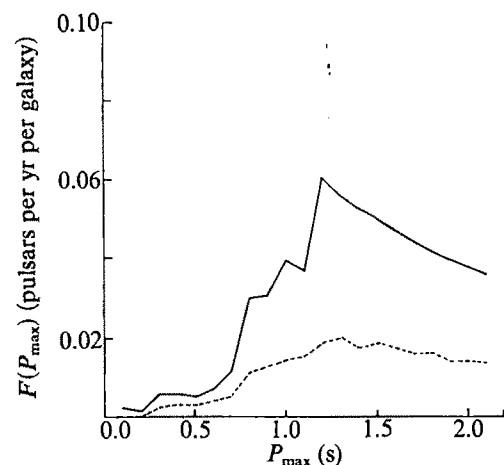


Fig. 1 The solid line shows $F_{\text{mean}}(P_{\max})$ while the dashed line shows $F_{951b}(P_{\max})$. We have selected the values at $P_{\max} = 1.3$ s to estimate pulsar birth rate.

$6.9^{+1.8}_{-1.2}$ yr. From a statistical comparison of this result with that given in equation (6) we can state with greater than 95% confidence that selection effects are significant and cannot be neglected in birth rate calculations. This throws all earlier calculations into doubt.

The present estimate of pulsar birth rate reduces but does not solve a fundamental problem on the origin of pulsars¹. The birth rate of supernovae, commonly believed to be the progenitors of pulsars, has lagged far behind pulsar birth rates, typical estimates ranging from 1 supernova per 50 yr to 1 per 150 yr (ref. 1). Our pulsar birth rate in equation (6) is significantly lower than earlier estimates, which are typically 1 pulsar per 7 yr but has not completely bridged the gap.

After submitting this work we received a preprint by E. S. Phinney and R. D. Blandford in which the concept of a current in the PP diagram has been used to test pulsar evolutionary models.

We thank J. H. Taylor for his comments and, in particular, for pointing out the importance of selection effects. We also thank R. Nityananda for helpful discussions.

Received 14 November 1980; accepted 27 February 1981.

1. Manchester, R. N. & Taylor, J. H. *Pulsars* (Freeman, San Francisco, 1977).
2. Davies, J. G., Lyne, A. G. & Seiradakis, J. H. *Mon. Not. R. astr. Soc.* **179**, 635–650 (1977).
3. Hulse, R. A. & Taylor, J. H. *Astrophys. J.* **201**, L55–59 (1975).
4. Manchester, R. N. *et al.* *Mon. Not. R. astr. Soc.* **185**, 409–421 (1978).
5. Lyne, A. G. *IAU Symp.* No. 95 (1980).
6. Ostriker, J. P. & Gunn, J. E. *Astrophys. J.* **157**, 1395–1417 (1969).
7. Narayan, R. & Vivekanand, M. (in preparation).
8. Lyne, A. G., Ritchings, R. T. & Smith, F. G. *Mon. Not. R. astr. Soc.* **171**, 579–597 (1975).

Spin-down episode of the X-ray pulsar Vela X-1

F. Nagase, S. Hayakawa, H. Kunieda, F. Makino, K. Masai & Y. Tawara

Department of Astrophysics, Nagoya University, Chikusa-ku, Nagoya, Japan

H. Inoue, K. Koyama, K. Makishima, M. Matsuoka, T. Murakami, M. Oda, Y. Ogawara, T. Ohashi, N. Shibasaki & Y. Tanaka

Institute of Space and Aeronautical Science, University of Tokyo, Meguro-ku, Tokyo, Japan

I. Kondo

Institute for Cosmic Ray Research, University of Tokyo, Tanashi, Tokyo, Japan

S. Miyamoto, H. Tsunemi & K. Yamashita

Department of Physics, Osaka University, Toyonaka, Osaka, Japan

Since the discovery¹ of a 283-s pulsation of an X-ray binary Vela X-1, several observations^{2–8} between 1975 and 1978 have revealed that its pulse period P was decreasing with an average value of $\dot{P}/P \approx -1.5 \times 10^{-4} \text{ yr}^{-1}$. The long-term spin-up which is considered to be caused by the angular torque of accreting matter, is a common feature of X-ray pulsars. The pulse periods of Vela X-1 were measured by the Japanese satellite Hakucho in March 1979 and March 1980, and the period in the latter was found to be significantly greater than in the earlier period despite the general spin-up between early 1976 and early 1979. Moreover, we have measured for the first time the spin-down rate $\dot{P}/P = (2.6 \pm 0.2) \times 10^{-3} \text{ yr}^{-1}$ in March 1980. This spin-down phenomenon is of great importance, as it is considered to be a manifestation of the dynamical behaviour caused by the internal structure of a neutron star.

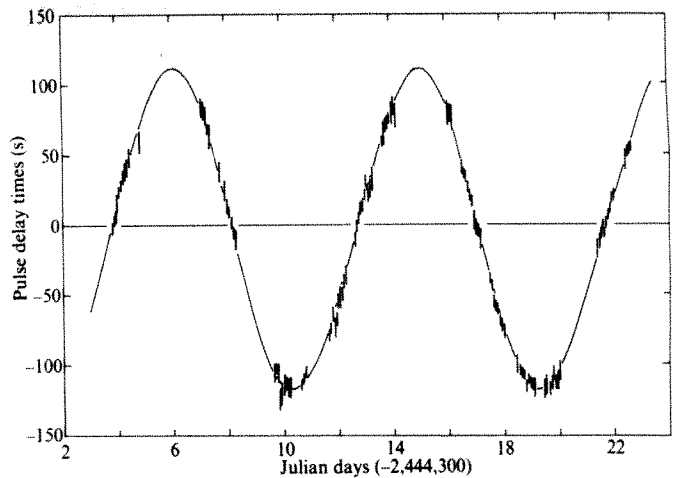


Fig. 1 Doppler delay curve of the Vela X-1 pulse arrival times observed in March 1980 compared with the calculated delay curve for the best-fit eccentric orbit with parameters given in Table 1. The vertical bars are measured delays with the length representing $\pm 1\sigma$ error limits.

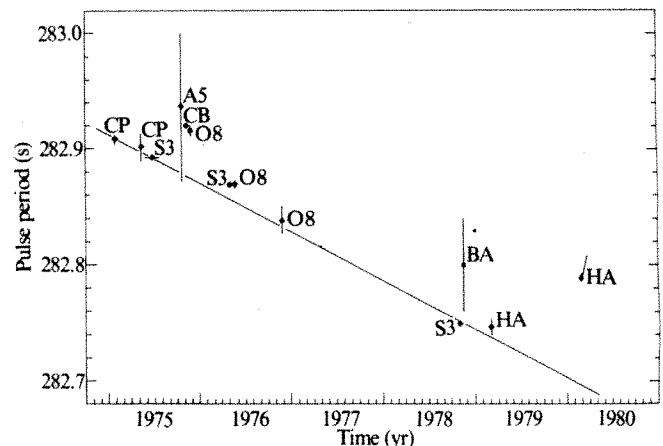


Fig. 2 The long-term pulse period history of Vela X-1. The measurements were made with Copernicus and Ariel V (CP and A5)⁶, SAS 3 (S3)^{2,4,7}, COS B (CB)³, OSO 8 (O8)⁵, Balloon (BA)⁸ and Hakucho (HA, present work). The solid line represents a long term spin-up at a rate of $\dot{P}/P = -1.5 \times 10^{-4} \text{ yr}^{-1}$ for guiding eyes. The short solid line attached to the 1980 point indicates the spin-down rate derived from the present observation (see text).

We observed Vela X-1 in 8.1–21.8 March (UT) 1979 and in 5.2–24.0 March (UT) 1980 from Hakucho, with its spin axis detectors in the energy range 1–22 keV. The binary period, the duration of eclipse, the pulse profile and its energy dependence are in good agreement with earlier results. The X-ray flux varied erratically by a factor of four in positive correlation with the softness ratio, which is the ratio of counting rate in the energy range 1–9 keV to that in range 9–22 keV, in addition to the variation associated with the binary phase. The X-ray luminosity observed in March 1980 was $(1-5) \times 10^{35} \text{ erg s}^{-1}$ in the energy range 2–22 keV at the estimated distance of 1.2 kpc, which is comparable with that in the previous observations.

Despite the differences in profile and intensity of each pulse the epoch of the pulse arrival time can always be identified by the sharp minimum of the pulse profile. Its epoch is obtained only within an accuracy of ~ 10 s for each pulse. Utilizing an uninterrupted pulse train consisting of several successive pulses for each satellite orbit improves the accuracy to 3 s. We obtained 35 and 115 epochs of pulses in 1979 and 1980, respectively.

Table 1 Best-fit orbital parameters for Vela X-1

Parameters	SAS 3† November 1978	Hakucho March 1979	Hakucho March 1980
$a_x \sin i (lt-s)$	112.3 ± 0.8	113.4 ± 3.8	114.4 ± 1.3
$f(M) (M_\odot)$	19.0 ± 0.4	19.5 ± 2.1	20.0 ± 0.7
e	0.094 ± 0.005	0.04 ± 0.07	0.091 ± 0.019
ω	$158^\circ \pm 5^\circ$	$182^\circ \pm 26^\circ$	$160^\circ \pm 7^\circ$
$P(s)$	282.7468 ± 0.0004	282.746 ± 0.006	282.7874 ± 0.0014
$\dot{P}$	$(-9 \pm 18) \times 10^{-9} \ddagger$	$(-20 \pm 10) \times 10^{-9}$	$(23.0 \pm 1.8) \times 10^{-9}$
$\tau (JD - 2,440,000)$	$3,823.53 \pm 0.13$	$3,938.4 \pm 0.7$	$4,307.68 \pm 0.18$
$P_{orb} (days)$	$8.9643 \S$	$8.965 \S$	$8.965 \S$

*All quoted uncertainties are single parameter 1σ confidence limits.

†After Table 1 of ref. 7

‡Corresponding to $\dot{P}/P = (-0.001 \pm 0.002) \text{ yr}^{-1}$.

§Orbital periods are fixed in the fitting.

The epoch of the n th pulse, after correcting for the heliocentric motion of the Earth, is expressed as

$$t_n = t_0 + P_0 n + \frac{1}{2} P_0 \dot{P} n^2 + a_x \sin i F(\theta, \omega, e, \tau)$$

where P_0 is the pulse period at a specified epoch t_0 , $\dot{P}$ is the rate of change of the period, $a_x \sin i$ is the projected semi-major axis of the orbit of the X-ray star, and F represents the effect of an eccentric motion with the periastron longitude ω , the eccentricity e , the time of periastron passage τ , and the mean anomaly $\theta = 2\pi (t - \tau)/P_{orb}$. The orbital period derived from X-ray eclipses between November 1978⁷ and March 1980 is $P_{orb} = 8.965 \pm 0.001$ days. This is considered a fixed parameter for the present analysis. Table 1 shows the orbital parameters obtained by fitting the observed pulse arrival times to the above equation. The best-fit Doppler curve for the data of March 1980 is shown in Fig. 1. All the parameters except P and $\dot{P}$ are consistent with earlier results^{2,3,7}.

Both the increase of the pulse period P between March 1979 and March 1980 and the positive value of $\dot{P}$ in March 1980 are remarkable. In March 1979 $P = 282.745 \pm 0.006$ (1σ)s, which lies nearly on the extrapolated line with $\dot{P}/P = -1.5 \times 10^{-4} \text{ yr}^{-1}$, whereas $P = 282.7874 \pm 0.0014$ (1σ)s in March 1980 which is significantly above this line, as shown in Fig. 2. This appreciable change in P seems to indicate a spin-down of the pulsar within a year. A similar increase of P was observed in late 1975³ and is suspected to be associated with a flare occurring in August 1975⁹. The increase in period observed in March 1980 may also be related to a flare in July 1979 observed by Hakucho¹⁰. Utilizing the data over two orbital cycles in March 1980 we obtain $\dot{P} = (23.0 \pm 1.8) \times 10^{-9}$, or $\dot{P}/P = (2.6 \pm 0.2) \times 10^{-3} \text{ yr}^{-1}$. Thus Vela X-1 occasionally undergoes a spin-down in the course of spin-up over a period of many years.

The observed increase of spin period can hardly be explained by fluctuations in the accretion rate^{11,12}, because the X-ray intensity in March 1980 is essentially the same as it was previously, although the model gives a favourable explanation for the period changes of the fast rotator such as Her X-1 and Cen X-3. Other external mechanisms of torque variation, such as mass ejection from the neutron star or the reversal of accreting stream in the wind-fed accretion, seem to have difficulty in accounting for the observed high rate of change of the pulse period. The internal torque acting on the crust of the neutron star, such as the braking stress of the rotation of crust caused by a lagging fluid core¹³, seems to be a possible explanation of the observed spin-down episode. A detailed consideration of the theoretical treatment of such a problem as well as further observations of the period change will be reported elsewhere.

Received 11 December 1980; accepted 24 February 1981.

- McClintock, J. E. *et al. Astrophys. J. Lett.* **206**, L99 (1976).
- Rappaport, S., Joss, P. C. & McClintock, J. E. *Astrophys. J. Lett.* **206**, L103 (1976).
- Ögelman, H. *et al. Astr. Astrophys.* **58**, 385 (1977).
- Rappaport, S. & Joss, P. C. *Nature* **266**, 683 (1977).
- Becker, R. H. *et al. Astrophys. J.* **221**, 912 (1978).
- Charles, P. A. *et al. Mon. Not. R. astr. Soc.* **183**, 813 (1978).

- Rappaport, S., Joss, P. C. & Stothers, R. *Astrophys. J.* **235**, 570 (1980).
- Staubert, R. *et al. Astrophys. J.* **239**, 1010 (1980).
- Watson, M. G. & Griffiths, R. E. *Mon. Not. R. astr. Soc.* **178**, 513 (1977).
- Tsunemi, H. *et al. Publ. astr. Soc. Jap.* (in the press).
- Ghosh, P. & Lamb, F. K. *Astrophys. J. Lett.* **223**, L83 (1978).
- Ghosh, P. & Lamb, F. K. *Astrophys. J.* **234**, 296 (1979).
- Baym, G., Lamb, D. Q., & Lamb, F. K. *Astrophys. J.* **208**, 829 (1976).

The very soft X-ray transient H1853+60

Robert A. Stern, Prahlad C. Agrawal*
& Guenter R. Riegler

Jet Propulsion Laboratory, California Institute of Technology,
Pasadena, California 91103, USA

The increased sensitivity, sky coverage and spectral range of X-ray observatories have revealed a wide range of transient phenomena. The 'classical' X-ray transients flare rapidly, sometimes increasing by more than an order of magnitude in X-ray flux in a matter of days, and decay on time scales of months. Such transients are mainly located in the galactic plane, often associated with 'steady' X-ray sources¹. Some high galactic latitude transients, most of which are optically unidentified, have been observed to vary on time scales of tens of seconds to days^{2,3}. Most transient sources have been detected at energies greater than 1 keV; more recently, however, softer X-ray transients have been detected with observations on HEAO 1 and other satellites⁴⁻⁹. We now report detection of a new soft X-ray transient, H1853+60, using data from the low energy detectors (LEDs) on the HEAO 1 A-2 experiment.

H1853+60 was observed by both LEDs 1 and 2 during a series of scans in ecliptic latitude on 13 May 1978 (see ref. 8 for details of the A-2 experiment operation). The source was detected at levels of 5.3 and 6.3 σ above background on two scans by LED 1 in the 0.16–0.4 keV band. There is no evidence for any emission above 0.5 keV. The source detection in LED 2 at the same sky location as LED 1 is significant, since that detector is offset by 6° from LED 1 in the scan plane, providing convincing evidence that the transient was not due to time-dependent particle events or instrumental effects. The source is unresolved by the 1.55° FWHM field of view of LED 1 along the scan direction.

Figure 1 plots the light curve of H1853+60 in the 0.16–0.4 keV band for 30 April–23 May 1978 (days 120–143). The source intensity has been corrected for the LED 1 collimator response (2.95° FWHM) perpendicular to the scan direction, assuming that the true source location is at the centre of the 90% confidence error box shown in Fig. 2. Scans with angular offsets $\geq 1.5^\circ$ from scan plane passing through the error box centre were not included so as to minimize possible systematic errors. Before 10 May, there is no evidence for a steady source (2σ upper limit ≈ 1 c.p.s.); an upper limit to flare-like behaviour is ≤ 3 c.p.s. From 11 to 16 May, the data are inconsistent (probability $\leq 1\%$) with any single count rate, even if we remove the two $\geq 5\sigma$ 'flare' scans. Summing scans together on a day-to-day basis, we also find no evidence for the presence of a source ($\leq 2\sigma$) before 10 May, but there is some ($\sim 3.5\sigma$) indication of post-outburst emission on 14 May; by 15 May the source is gone.

The peak LED 1 count rate on 13 May was 9.3 ± 1.5 c.p.s. above background, corresponding to a broad-band flux of $\sim 3 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the 0.16–0.4 keV band. H1853+60 was not detected in a steady or flaring state 6 months earlier when HEAO 1 scanned the same area of the sky. In addition, no steady source was detected in the HEAO 1 MED or HEDs (2–10 keV band)¹⁰, and an examination of individual MED and

* Permanent address: Tata Institute of Fundamental Research, Bombay, India.

HED scans while H1853+60 was flaring shows no evidence for X-ray emission (F. Marshall, personal communication).

Because of the relatively poor spectral resolution of proportional counters at such low energies, we can only obtain upper limits to the source temperature and interstellar medium column density of $T \leq 10^6$ K and $N_H \leq 10^{20.6} \text{ cm}^{-2}$, respectively. The spectrum does, however, exhibit a steep rise at the lowest detectable energies, suggesting that much of the energy output of the transient event was below our detection threshold of 0.16 keV.

A search for possible optical counterparts for H1853+60 in the available literature has produced the following candidates: the Mira variable AD Dra which is located at the northeastern corner of the error box, the faint (15th magnitude) galaxy MCG+10-27-001, the centre of the Zwicky cluster ZC1851.3+6032, the nearby double star Gliese 725A and B, and a number of 8th and 9th mag SAO stars. There are no previously discovered X-ray sources, Yale catalogue bright stars, RS CVn systems and other variable stars, white dwarfs, or other proper motion or nearby stars in the error box.

Gliese 725AB (dM4, dM4.5; HD173739) has been reported as a flare star¹¹. Although there is no evidence for hydrogen emission—a characteristic usually associated with UV Ceti flare stars—Ca II emission has been observed (this, however, is common in late-type dwarfs)¹². Few X-ray flares from UV Ceti stars have been observed, and those that have been are generally seen at harder (≥ 2 keV) energies than H1853+60 (refs 13, 14). AT Mic, which was detected by both the HEAO 1 LEDs and MED, had a flaring spectrum characterized by $T \geq 10^7$ K (ref. 14). The detection of H1853+60 at $\geq 5\sigma$ for two successive scans separated by 0.1 day is also not consistent with typical UV Ceti flare risetimes¹³. If HD173739 proves to be the optical counterpart of H1853+60, one possibility is that we have observed a 'long-duration enhancement' such as that seen in the solar corona, with risetimes of the order of hours, and decay times of a day or more; even these enhancements, however, have characteristic temperatures only slightly lower ($\sim 3\text{--}7 \times 10^6$ K) than those of a solar flare^{15,16}.

Although the stellar flare hypothesis cannot be definitely ruled out, we must also consider the similarity of H1853+60's spectral and temporal characteristics to several other transients observed by the HEAO 1 LEDs^{8,9}, suggesting a common emission mechanism for these events. As the transients referred to were initially discovered using 1-day superpositions of the LED data, which have been examined extensively (J. Nousek, personal communication), it is unlikely that many more transients as strong as these will be discovered, except on an individual scan basis. The only unusual class of optical object common to the transients' error boxes are Mira variables, of which $\sim 4,500$ are known. Given an error box area of $\sim 1.5 \text{ deg}^2$,

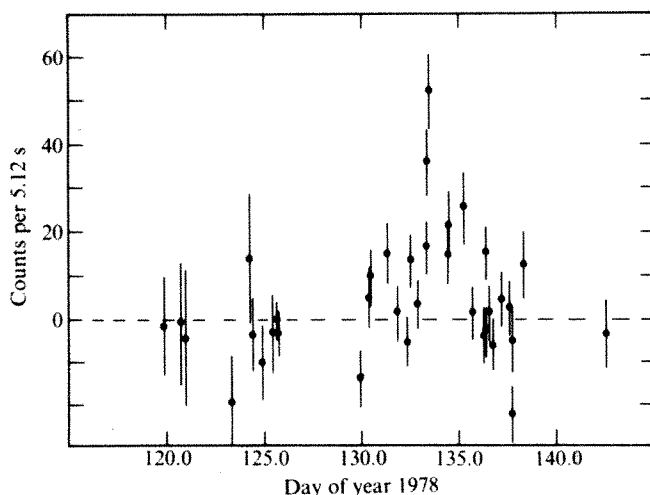


Fig. 1 Light curve of H1853+60 in the 0.16–0.4 keV band. Error bars are $\pm 1\sigma$.

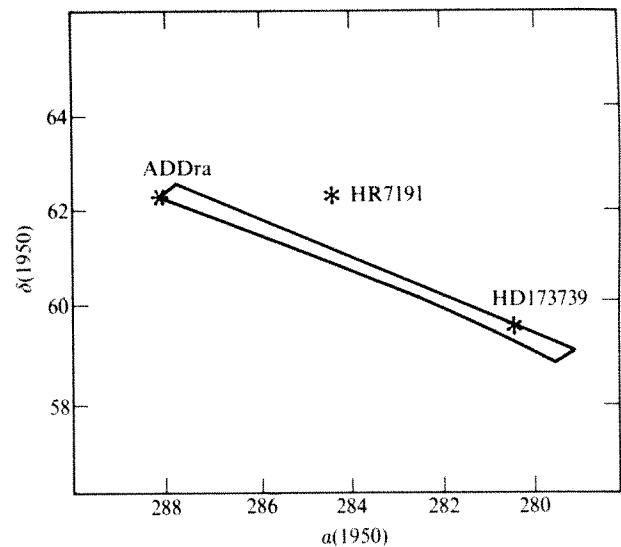


Fig. 2 90% confidence error box for H1853+60; HD173739 = GL725A.

there is $\geq 10\%$ probability that a Mira variable would be found within a typical LED error box by chance alone. (Ref. 8 reviews the possibility of such transient X-ray emission in supernova outbursts, undiscovered RS CVn systems and shock heating of the atmospheres of the Mira variables described above; none of these mechanisms is likely to account for emission in H1853+60.)

The suggestion of a U Gem-like outburst from an as yet undiscovered dwarf nova (also discussed in ref. 8) is intriguing because of the similarity of the spectra observed in both U Gem⁵ and H1853+60, and the approximately equal risetimes for the outbursts. Although U Gem was also observed in harder (2–10 keV) X rays¹⁷, the 0.16–0.4 keV peak flux from H1853+60 is about 10 times less than that from U Gem, and thus if the harder X rays scale proportionately, the hard X-ray flux from such a dwarf nova outburst would fall below the sensitivity level of the MED or HEDs for single scans ($\sim 4\text{--}6 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$). If we assume that H1853+60 is indeed a dwarf nova with the same X-ray luminosity as U Gem, we derive a distance of about three times that of U Gem, or ~ 300 pc, based on usual distance estimates⁵. If the optical outburst is scaled in luminosity, the optical counterpart of H1853+60 would be at $V \sim 12$ at outburst and $V \sim 16$ when quiescent. Although fainter U Gem-type systems have been discovered¹⁸, the completeness of the sample of U Gem stars down to a particular limiting magnitude is unknown; further optical observations will be needed to confirm the dwarf nova hypothesis.

We acknowledge useful discussions with J. Nousek. This work was performed at the Jet Propulsion Laboratory under NASA contract NAS 7-100. P.C.A. is a NAS-NRC Senior Research Fellow.

Received 3 November 1980; accepted 10 February 1981.

- Willmore, A. P. *Proc. R. Soc. A* **350**, 463 (1976).
- Johnston, M. D., Griffiths, R. E. & Ward, M. J. *Nature* **285**, 26 (1980).
- Schrijver, J. *et al. Astr. Astrophys.* **69**, L1 (1978).
- Rappaport, S., Cash, W., Doxsey, R., McClintock, J. & Moore, G. *Astrophys. J. Lett.* **187**, L5 (1974).
- Mason, K. O., Lampton, M., Charles, P. & Bowyer, S. *Astrophys. J. Lett.* **226**, L129 (1978).
- Heise, J. *et al. Astrophys. J. Lett.* **202**, L73 (1975).
- Kahn, S. M. *et al. Astrophys. J. Lett.* **234**, L107 (1979).
- Nousek, J., Cardova, F. & Garmire, G. *Astrophys. J.* (in the press).
- Jensen, K. & Nousek, J. *Bull. Am. astr. Soc.* **12**, 820 (1980).
- Marshall, F. E. *et al. Astrophys. J. Suppl.* **40**, 657 (1979).
- Pettersen, B. R. *Inst. Theor. Astrophys., Blindern-Oslo Rep.* no. 46 (1976).
- Joy, A. H. & Abt, H. A. *Astrophys. J. Suppl.* **28**, 1 (1974).
- Gorenstein, P. & Tucker, W. H. A. *Rev. Astr. Astrophys.* **14**, 373 (1976).
- Kahn, S. M. *et al. Astrophys. J. Lett.* **234**, L107 (1979).
- Underwood, J. H. *et al. Prog. Aeron. Astr.* **48**, 179 (1976).
- McCombie, W. J. & Rust, D. M. *Sol. Phys.* **61**, 69 (1979).
- Swank, J. H., Boldt, E. A., Holt, S. S., Rothschild, R. E. & Serlemitsos, P. J. *Astrophys. J. Lett.* **226**, L133 (1978).
- Kukarkin, B. V. *et al. General Catalog of Variable Stars and supplements* (Nauka, Moscow, 1969–74).

The Kelvin equation and the capillary condensation of water

L. R. Fisher, R. A. Gamble & J. Middlehurst

CSIRO Division of Food Research, PO Box 52, North Ryde, New South Wales 2113, Australia

The Kelvin equation¹ relates the equilibrium vapour pressure of a liquid to the curvature of the liquid-vapour interface. It predicts that undersaturated vapours will condense in channels of sufficiently small dimensions. While the applicability of the Kelvin equation to organic liquids with meniscus radii as low as 4 nm has been verified by direct experiment², its applicability to water for meniscus radii below 4 nm is unverified, and has been doubted^{3,4}. We have used Fizeau interferometry to measure directly the capillary condensation of water in a wedge between fused silica surfaces. Measurements were made at relative vapour pressures (P/P_s) from 0.996 down to 0.945, corresponding to theoretical meniscus radii from ~120 to 9 nm. We find only a small discrepancy between our experimental results and the Kelvin equation in the vapour pressure range covered, provided that correct account is taken of the effect of the adsorbed film, up to 200 nm thick, on the meniscus shape.

The Kelvin equation forms the basis of critical nucleation theory⁵, and is widely used in the interpretation of such diverse phenomena as adhesion⁶, the enhanced solubility of small particles⁷ and the retention and transport of liquids in porous materials⁸⁻¹⁰. Because water is an important liquid in the above applications, we need direct experimental evidence concerning its capillary condensation behaviour.

Cross and Picknett¹¹ have shown experimentally that water obeys the Kelvin equation for meniscus radii $> \sim 4 \mu\text{m}$, but most applications of the equation involve much smaller radii. Other experiments^{3,4} have purported to show that the Kelvin equation grossly underestimates the degree of capillary condensation of water, but these results are now generally ascribed to the presence of impurities^{4,12}, and there have been no reliable studies in the important region covering menisci with radii $< 1 \mu\text{m}$.

For the present experiments, two silica surfaces (either two optical flats held as a very narrow wedge, or a spherical surface of radius 69 m on a flat surface) were clamped between Teflon supports in an evacuated isothermal enclosure. The vapour pressure in the enclosure was controlled by an aqueous potassium nitrate solution. Stringent cleaning procedures, including prolonged boiling in nitric acid followed by prolonged boiling in freshly distilled water and finally blowing dry with filtered dry nitrogen, were applied to the surfaces.

Each silica surface was first tested for wettability by exposing it to the vapour above warm distilled water and observing the appearance of coloured interference fringes on the surface (the so-called 'steam test')¹³. Surfaces for which the contact angle is non-zero exhibit hazy condensation due to microdroplets, rather than clear, broad fringes. At the end of an experimental run, the silica surfaces were carefully separated so that the condensed liquid collected at one point, where it evaporated. The surfaces were then examined in a strong light to ensure the absence of any residue. This test is sufficiently sensitive to detect dissolved material equivalent to 0.01 of a monolayer. The surfaces were then retested for wettability, and finally atomic absorption spectroscopy was used to test for the presence of inorganic ions (for example, K^+) in a small volume (0.1–0.2 ml) of water rinsed over each surface after each experiment. Only if all these tests were passed was the result of an experiment accepted. Also, no experimental run was accepted unless the same results were obtained with increasing and decreasing relative vapour pressure.

The appearance of capillary-condensed water in a wedge between two optical flats is shown in Fig. 1. The Fizeau fringes are shifted with respect to their position when the surfaces were dry because of the presence of the adsorbed water film. This shift was used to calculate the adsorbed film thickness Z_1 (Table 1). The distance D_{men} between the silica surfaces at the meniscus can then be calculated from the fringe positions in the dry state. The results of condensation experiments for both the wedge and the sphere-on-flat configuration are given in Fig. 2, in which $-\ln(P/P_s)$ is plotted against $1/D_{\text{men}}$.

The finding that water reversibly forms relatively thick adsorbed films on a fused quartz surface agrees with the results of Pashley and Kitchener's¹³ ellipsometric study. These results clearly have a bearing on the question of water 'structure' at solid interfaces^{4,12,13}, but we cannot add any comment to that already made by Pashley and Kitchener who have shown that very thick films are consistent with calculations from double layer repulsion theory. However, several authors have postulated deviations from the Kelvin equation because of the influence of such a thick adsorbed film (more correctly, the influence of a long-range adsorption potential) on meniscus curvature¹⁴⁻¹⁷, and it is with this aspect that we are now concerned.

All authors find the same form for the Kelvin equation for the case of capillary condensation in a slit-shaped pore^{17,18}:

$$D_{\text{men}} = \frac{2\gamma V_m}{R_s T \ln(P/P_s)} + 2Z_1 - \left\{ 2/R_s T \ln(P/P_s) \right\} \int_{Z_1}^{D_{\text{men}}} F(Z) dZ \quad (1)$$

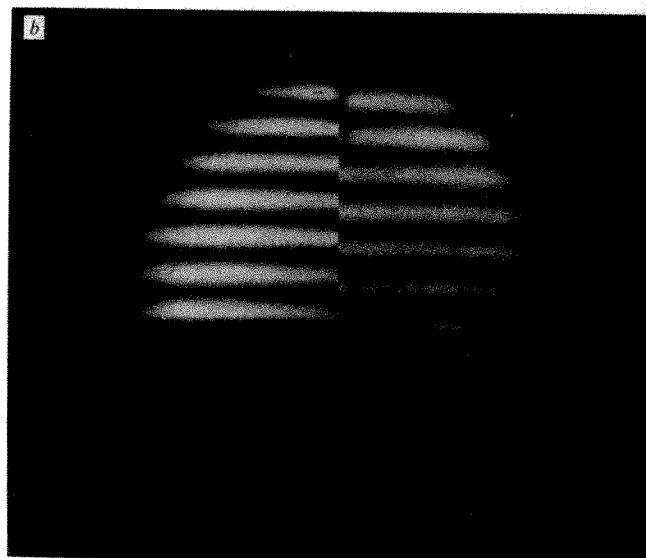
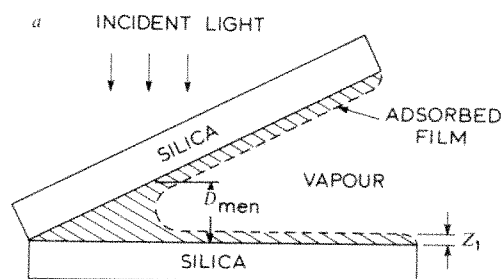


Fig. 1 a, Diagrammatic representation of liquid condensed in wedge between two optical flats. b, Comparison of appearance of wedge, viewed normally in monochromatic light ($\lambda = 546.1 \text{ nm}$) before (right-hand side) and after (left-hand side) exposure to water vapour at $P/P_s = 0.9969$. Note that all fringes have shifted equally, showing that the Laplace pressure in the condensed liquid does not distort the silica surfaces. Silica surfaces are in contact at the bottom of the picture, the bright mark at the top is a gold foil spacer.

Table 1 Experimental adsorbed film thickness Z_1 of water condensed in a wedge between flat silica surfaces

Relative vapour pressure P/P_s	Z_1 (nm)	$-0.43/\ln(P/P_s)$ (nm)
0.90	3*	4.1
0.95	5*	8.9
0.965	8*	12.0
0.98	20, 27*	21
0.9925	(59 ± 5)†	57
0.9944	(82 ± 5)†	77
0.9955	(120 ± 10)†	95
0.9969	(200 ± 30)†	139
1.00 ± 0.001	150*‡	∞

* Data from ref. 13: we obtained agreement with these figures in an experiment where the condensation was reversible, but the results were rejected because of trace quantities of residue at the end of the run.

† Present work.

‡ Agrees with calculations based on double-layer theory.

where γ is the surface tension, V_m the molar volume, P' is the vapour pressure at which the liquid just irreversibly evaporates, Z the normal distance from the solid surface, R_0 the gas constant and T the absolute temperature (this form of the equation neglects corrections which may be necessary at extremely high meniscus curvatures¹⁹). The first term is the simple Kelvin term, the second makes allowance for the adsorbed film by simple subtraction, and the third is a correction factor to allow for the effect of adsorption potential on meniscus curvature. The function $F(Z)$ is the adsorption isotherm of water on a flat surface. The accuracy of the adsorption data does not warrant numerical evaluation of this integral over the whole meniscus, but fortunately the simple expression Z_1 (nm) = $0.43/\ln(P_s/P)$ follows the data sufficiently closely over much of the range of interest, so $F(Z)$ can be put in the form $F(Z) = 0.43R_0T/Z$, the integral of which can be obtained by a simple iterative method¹⁷.

The uncorrected Kelvin equation, the Kelvin equation with allowance for the adsorbed film by simple subtraction, and equation (1) are plotted in Fig. 2. Only equation (1) is in reasonable agreement with the experimental data, although there is still a significant discrepancy. This discrepancy may well be due to the breakdown of the 'boundary layer approximation' implicit in equation (1). That is, it may not be possible to treat

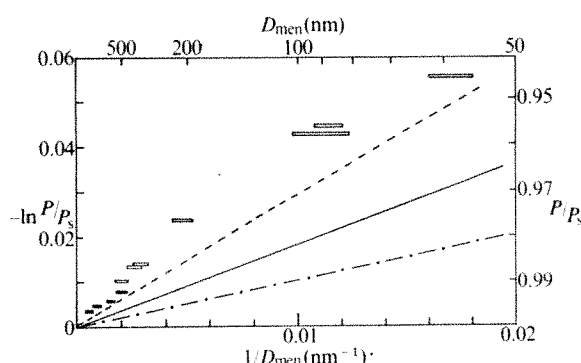


Fig. 2 Experimental results for capillary condensation of water in a wedge between silica optical flats (closed bars) and in the gap between a spherical silica surface and contacting silica optical flat (open bars). The size of the rectangles represents standard errors, based on at least three measurements of the position of each dark fringe. —, Uncorrected Kelvin equation, first term of equation (1). —, Kelvin equation with correction by subtraction for thickness of adsorbed film, first two terms of equation (1). - · - · - ·, Fully corrected Kelvin equation (1). ■, Experimental results from wedge configuration. □, Experimental results from sphere-on-flat configuration. For the sphere-on-flat experiments, values of Z_1 were calculated from Z_1 (nm) = $-0.43/\ln(P/P_s)$.

the adsorption and capillary potentials as strictly additive when Z_1 is of the order of $D_{men}/2$. This point has been discussed in detail elsewhere²⁰. The fact that the difference between theory and experiment apparently decreases with increasing D_{men} supports this interpretation.

We conclude that water reversibly forms relatively thick adsorbed films (up to 200 nm) on fused silica surfaces. The origin of the stability of these films is not yet fully understood¹³. The presence of these films affects the meniscus shape for capillary-condensed water, and this effect accounts almost entirely for the capillary condensation behaviour of water. The small residual difference between theory and experiment may well be a consequence of the breakdown of the 'boundary-layer' approximation used in current theories. A fuller report of this work will be given elsewhere.

Received 10 November 1980; accepted 6 March 1981.

1. Thomson, W. *Proc. R. Soc. Edinburgh* **7**, 63–68, (1870); **42**, 448–452 (1871).
2. Fisher, L. R. & Israelachvili, J. N. *Nature* **277**, 548–549 (1979); *J. Colloid Interface Sci.* (in the press).
3. Skinner, L. M. & Sambles, J. R. *Aerosol Sci.* **3**, 199–210 (1972).
4. Clifford, J. in *Water: A Comprehensive Treatise* Vol. 5 (ed. Franks, F.), 75–132 (Plenum, New York 1975).
5. Becker, R. & Döring, W. *Ann. Phys.* **24**, 719–754 (1935).
6. Zimon, A. D. *Adhesion of Dust and Powder*, 81–90 (Plenum, New York, 1969).
7. Kahlweit, M. *Adv. Colloid Interface Sci.* **5**, 1–35 (1975).
8. Dullien, F. A. L. & Batra, V. K. *Ind. Eng. Chem.* **62**, 25–53 (1970).
9. Everett, D. H. & Haynes, J. M. *Colloid Sci.* **1**, 123–172 (1973).
10. Boucher, E. A. *J. mater. Sci.* **11**, 1734–1750 (1976).
11. Cross, N. L. & Picknett, R. G. *Trans. Faraday Soc.* **59**, 846–855 (1963).
12. *J. Colloid Interface Sci.* **36**, 415 (1971).
13. Pashley, R. M. & Kitchener, J. A. *Colloid Interface Sci.* **71**, 491–500 (1979).
14. Broekhoff, J. C. P. & de Boer, J. H. *J. Catal.* **9**, 8–14 (1967).
15. Deryagin, B. V. *Proc. 2nd Int. Congr. Surface Activity* **2**, 153–159 (1957).
16. Deryagin, B. V. & Churaev, N. V. *Chem. Phys.* **70**, 596–597 (1979).
17. Philip, J. R. *J. chem. Phys.* **66**, 5069–5075, (1977); **70**, 598 (1979).
18. Broekhoff, J. C. P. & de Boer, J. H. *J. Catal.* **10**, 391–400 (1968).
19. Buff, F. P. in *Handbuch der Physik* Vol. 10 (ed. Flügge, S.) 281–304 (Springer, Berlin, 1960).
20. Philip, J. R. *J. chem. Phys.* **67**, 1732–1741 (1977).
21. Derjaguin, B. V. & Churaev, N. V. *J. Colloid Interface Sci.* **49**, 249–255 (1974).

Hydrogen production from photolysis of steam adsorbed onto platinized SrTiO₃

R. G. Carr & G. A. Somorjai

Materials and Molecular Research Division, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, California 94720, USA

Photodissociation of high pressure (1–8 atm) water vapour adsorbed onto the partially platinized surface of single crystal SrTiO₃ has been observed. The reaction is catalytic; several thousands of monolayers of hydrogen have been produced. The reaction requires light of greater than bandgap (3.1 eV) energy, and proceeds vigorously only if elevated temperatures (375–450 K) and high pressures are used. Recombination of products and poisoning of the metal apparently diminishes the reaction rate after several hours. Previous work on this reaction has been done in the presence of an electrolyte, generally at room temperature. The present work is a significant departure from such an environment, and is important because the surface is not contaminated with an electrolyte. The surface is therefore accessible to ultrahigh vacuum surface analytical techniques such as electron spectroscopy, before and after the reaction. From these studies detailed information about the mechanism may be obtained. In addition, the existence of a surface hydrogen species at elevated temperatures allows us to combine water photolysis directly with other catalytic reactions, such as ammonia synthesis with N₂ and methane synthesis with CO or CO₂.

A reaction chamber made of stainless steel was fitted with a sapphire window and placed in the oven of a gas chromatograph. Light from a high-pressure mercury discharge lamp passed through the window and was incident on the surface of the crystal. The intensity was $\sim 10^{16}$ photons $\text{cm}^{-2} \text{s}^{-1}$ in the energy range above the band gap. A single crystal wafer of SrTiO_3 with nominal (111) orientation was rough-polished and coated *in vacuo* with platinum evaporated from a tungsten filament. We used non-conducting, stoichiometric SrTiO_3 as bulk electrical conductivity has little influence on localized surface reactions. The roughness of the surface ensured that it was only partially covered with platinum. The photoreaction rate was highest with a transparent coating of platinum and decreased with heavier, more opaque coatings, which would cover all semiconductor surface sites.

Ultrapure water was admitted into the evacuated cell at the beginning of an experiment, and pumped through a loop which included the cell and an automatic sampling valve, so that gas samples could be withdrawn periodically. A sample was then analysed in the gas chromatograph, which was fitted with a thermal conductivity detector. With argon carrier gas, we could detect 10 monolayers of H_2 easily, and O_2 with about 20 times less sensitivity. The whole system accessible to water vapour was kept at uniform temperature so that high pressures could be achieved. Figure 1 shows the cell design.

Figure 2 shows H_2 accumulation curves as a function of time for different temperatures, where the saturation vapour pressure of water was used in each case. To inhibit the recombination of H_2 and O_2 products, a few grammes of 0.025 mm Cu foil was placed in the reaction cell to getter the O_2 present. Without this extra copper to oxidize, and thus remove active O_2 from the system, less than half the H_2 would be detected. Oxidation of Cu by water with a release of H_2 was ruled out by blank experiments. Thus the recombination of hydrogen and oxygen is a significant process in these conditions. This is not surprising in view of the presence of platinum and a light activated semiconductor, either of which might act as a catalyst. Because of this back reaction, O_2 was not detected. Blank experiments showed no hydrogen production without platinum on the crystal, nor with platinum on an alumina substrate, and very poor production when platinum was deposited on the dark side of the crystal.

The curves in Fig. 2 show an induction period which depends on exposure to water vapour. If the crystal is exposed to water vapour in the dark for 1–2 h, hydrogen production will proceed immediately on illumination. However, when the crystal is illuminated *in vacuo* for the same period and then exposed to water vapour, the induction period is observed. Thus there is

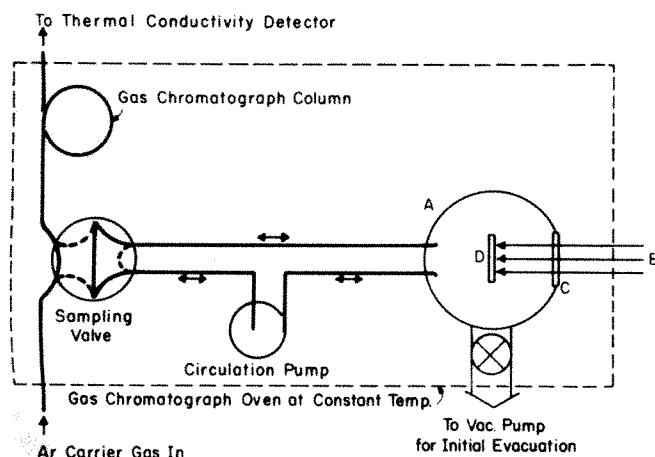


Fig. 1 Apparatus: A, stainless steel cell; B, UV light from Hg lamp; C, sapphire window; D, SrTiO_3 crystal with Pt on the side facing the light. Water vapour circulates through cell, pump, and sampling valve.

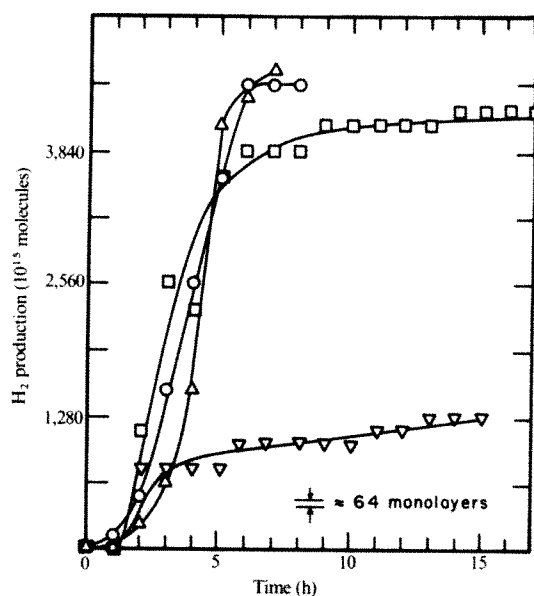


Fig. 2 H_2 product accumulation for four different temperatures at H_2O saturation vapour pressures. Δ , 160 °C, 6 atm; $\circ$, 140 °C, 3.5 atm; $\square$, 120 °C, 2 atm; ∇ , 100 °C, 1 atm.

evidence for a slow surface reaction, such as hydroxylation or displacement, which seems to be a necessary precursor for the efficient catalytic photolysis of water observed after the induction period. During efficient hydrogen production, a rate of as much as 500 monolayers per hour has been observed. Within ~ 10 h, however, the hydrogen production rate diminishes. Besides equilibration of the production and recombination rates there is also a surface poisoning effect which contributes to this decline. If a crystal used in the experiment is not regenerated by replacement of the platinum coating, it will yield much less hydrogen in a subsequent experiment. A similar deactivation was noted by Wrighton *et al.*¹ in the SrTiO_3 -Pt electrochemical cell.

It was possible to carry out the photolysis of water at less than saturation vapour pressure. At constant pressure, the reaction rate declines with increasing temperature, as expected, because the recombination over platinum is more facile at higher temperatures. At constant temperature, the rate of hydrogen production increases with increasing pressure.

Recently we reported the photocatalysed dissociation of water vapour on platinum coated and metal-free SrTiO_3 crystal surfaces which were covered with a film of NaOH (refs 2, 3). The need for NaOH was traced to the importance of hydroxylation of the oxide surface. In the conditions of the present experiment the surface is expected to be hydroxylated⁴, and this effect is apparently equivalent functionally to the effect of NaOH.

In conclusion, we have observed evidence for the following steps during photolysis of water on platinized SrTiO_3 : (1) generation of electron-hole pairs by light of greater than bandgap energy; (2) dissociative chemisorption of water vapour on the hydroxylated semiconductor surface; (3) diffusion of a surface hydrogen species to platinum islands and desorption of hydrogen from the islands; and (4) recombination of hydrogen and oxygen which limits the amount of hydrogen that can be generated in a batch reactor.

This work was supported by the Division of Chemical Sciences of the Office of Basic Energy Sciences of the US Department of Energy under contract W-7405-ENG-48.

Received 24 November 1980; accepted 18 February 1981.

1. Wrighton, M. S., Wolczanski, P. T. & Ellis, A. B. *J. Solid St. Chem.* **22**, 17 (1977).
2. Wagner, F. T. & Somorjai, G. A. *Nature* **285**, 559 (1980).
3. Wagner, F. T., Ferrer, S. & Somorjai, G. A. *J. Am. chem. Soc.* **102**, 5494 (1980).
4. Ferrer, S. & Somorjai, G. A. *Surf. Sci.* **94**, 41 (1980).

Fluid inclusion Rb–Sr isochrons for dating mineral deposits

T. J. Shepherd & D. P. F. Darbyshire

Isotope Geology Unit, Geochemistry and Petrology Division,
Institute of Geological Sciences, 64 Gray's Inn Road,
London WC1X 8NG, UK

The precise dating of mineral deposits has important implications for understanding the lifetime of ore-forming systems and their temporal relationship to various geological events. Not all deposits provide material suitable for conventional radiometric dating, or data which are amenable to unequivocal interpretation. Nor is it easy to verify that such samples have remained isotopically 'closed' since their formation. Fluid inclusions provide an alternative 'closed system' approach. By analysing the isotopic composition of Rb and Sr in inclusions within quartz from the Carrock Fell tungsten vein deposits, Cumbria, we demonstrate here the feasibility of obtaining Rb–Sr fluid inclusion isochrons. The age for the mineralization 392 ± 5 Myr is in good agreement with the mean K–Ar age 387 ± 6 Myr for the vein micas, and confirms the significance of fluid inclusions as precise Rb–Sr geochronometers.

Fluid inclusions are small, usually microscopic, volumes of fluid trapped within crystals during their growth from the fluid. To be used as geochronometers their origin must be known and also that nothing has been added or lost from the inclusion after sealing. The inclusions have usually been trapped at more than one stage in the growth and subsequent fracturing and healing of the crystal. For the purposes of dating, however, as long as the various inclusions were formed in a relatively short time interval, the crystal can be treated as a single temporal event. For a more complex chronology, samples representing individual episodes of mineral deposition can be selected.

In most geological conditions inclusions do not leak¹ although exceptions are excessive overheating² and the development of an intracrystalline deformation fabric³ within the host mineral. The former leads to thermal rupture with a complete emptying of the inclusion, whereas the latter may vary from a simple 'drawing-out' of the inclusion with minor leakage, to total rupture and destruction. Host minerals with little or no cleavage and good mechanical strength are thus preferred. Quartz, a ubiquitous mineral of hydrothermal ore deposits is suitable, particularly as its high chemical purity permits analysis of the inclusion contents independently of ion impurities in the host crystal.

Inclusions provide a means of applying basic geochronological concepts to situations where conventional mineral and rock studies have failed. Excluding fission track dating, the use of inclusions as geochronometers has the same restrictions as all radiometric methods of dating.

(1) The mineral or rock samples must have behaved as 'closed' systems with respect to the parent–daughter isotopes according to the equation

$$D = D_0 + P(e^{\lambda t} - 1)$$

where D is the total number of daughter atoms present in the system; D_0 the number of daughter atoms initially present; P , the number of parent atoms remaining in the system; λ , the decay constant for the parent radionuclide; and t , the time elapsed ('age' of the system). (Inclusions in transparent minerals can be checked before analysis to ascertain whether thermal

rupture or intracrystalline deformation has caused them to leak.)

(2) The mineral or rock samples were formed in a time interval that was short compared with their age. (Modern hydrothermal–hot spring systems driven by local igneous heat sources have lifetimes of a few thousand to approximately one million years. Presumably palaeohydrothermal systems associated with igneous rocks had similar lifetimes. Hydrothermal ores precipitated on the sea floor are thought to have had formation times of a few thousand years because they contain little detritus, even though the rocks above and below are normal detrital sediments⁴.)

(3) A value for the initial number of daughter atoms (D_0) can be assumed or calculated.

Inclusions in quartz can be thought of as droplets of fluid preserved, totally isolated from their external environment, and perfect for Rb–Sr dating. The only disadvantage is their small size, normally $<20 \mu\text{m}$ in diameter. For a fluid containing 1,000 p.p.m. Sr a single inclusion may provide less than 10^{-12} g Sr for analysis.

To test the feasibility of using fluid inclusions as Rb–Sr geochronometers for ore deposits we have taken samples from the tungsten deposits at Carrock Fell, Cumbria. These deposits are characterized by a series of steeply dipping north–south wolframite-bearing quartz veins developed at the contact of the Caledonian Skiddaw Granite with the Lower Palaeozoic Skiddaw Slate Series. Mineral paragenesis studies^{5,6} suggest a main period of hydrothermal activity characterized by continuous quartz deposition, with wolframite and apatite confined almost exclusively to the earliest generations of quartz. Much of the quartz is therefore economically 'barren'. Late-stage hydrothermal activity is represented by a network of cross-cutting east–west carbonate veins and stringers, some of which cement and infill vuggy parts of the quartz veins. Compared with the tin–tungsten deposits of south-west England, the mineralization constitutes a simpler hydrothermal event, which with data from a previous geochronological study⁵ was ideal for investigation. The samples were selected from an extensive suite of vein quartz material collected from underground workings at the Carrock Fell Mine and were representative of both the early and late phases of quartz deposition. The quartz contains abundant two-phase fluid inclusions consisting of a liquid phase and complementary vapour phase, the latter occupying $<25\%$ of the total inclusion volume. They average $5\text{--}30 \mu\text{m}$ in diameter, are ovoid or multifaceted in form, and occur in both isolated non-planar groups and discontinuous intersecting planar arrays. Genetically the inclusions are primary, but because it is difficult to relate them to obvious growth features they are best described as pseudosecondary^{7,8} (they correspond to fluids trapped along microfractures during crystal growth). Collectively the inclusions can be treated as the product of a single mineralizing

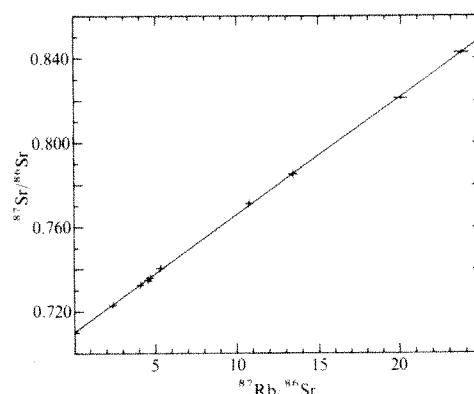


Fig. 1 Fluid inclusion Rb–Sr diagram. Age, 393 ± 5 Myr; intercept, 0.70997 ± 0.0004 ; MSWD, 1.42.

Table 1 Rb–Sr data for inclusion fluids, Carrock fell tungsten deposit, Cumbria

Sample no.	Rb (p.p.m.)	Sr (p.p.m.)	$^{87}\text{Rb}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}$
CF-79-10	446	287	4.511	0.73465
	512	281	5.289	0.74043
	475	298	4.625	0.73577
	388	278	4.041	0.73239
CF-79-21	1,225	179	20.01	0.82071
	1,477	183	23.69	0.84253
CF-79-22	505	164	8.971	0.75892
CF-79-30A	166	207	2.318	0.72296
CF-79-30B	765	166	13.46	0.78496
	1,112	301	10.77	0.77102

event. Unlike the pseudosecondaries, secondary inclusions are less well developed and clearly associated with the late-stage carbonate fluids. They can be recognized by their distinctive chemistry and behaviour during low temperature microthermometric studies. Samples suspected of containing either secondaries or inclusions which had leaked were rejected.

The sample preparation and mass spectrometric analysis will be described elsewhere. Conventional isotope dilution techniques⁹, modified to accommodate low level, small samples, were used (the results, given in Table 1, are plotted in Fig. 1). Analytical errors ($\pm 1\%$ $^{87}\text{Rb}/^{86}\text{Sr}$; $\pm 0.03\%$ $^{87}\text{Sr}/^{86}\text{Sr}$) agree with those normally assigned to precision Rb–Sr mineral/whole rock isochrons. The Rb–Sr age was calculated¹⁰, using an ^{87}Rb decay constant of $1.42 \times 10^{-11} \text{ yr}^{-1}$, with errors quoted at the 2σ level. A MSWD of 1.42 indicates that the data satisfy the criteria for an isochron.

The fluid inclusion Rb–Sr age $392 \pm 5 \text{ Myr}$ is in good agreement with the mean K–Ar age for the hydrothermal vein muscovites $387 \pm 6 \text{ Myr}$ determined by Rundle⁵. Furthermore the age of the main Lake District batholith, of which the Skiddaw Granite is part, has been established as $\approx 395 \text{ Myr}$ (Shap Granite, Rb–Sr age $394 \pm 3 \text{ Myr}$ (ref. 11), K–Ar age $397 \pm 7 \text{ Myr}$; Skiddaw Granite, K–Ar ages $392 \pm 4 \text{ Myr}$ (ref. 5) and $399 \pm 6 \text{ Myr}$ (ref. 12)). As the veins cannot be older than the granite and there is no reason to doubt the K–Ar vein muscovite age, the agreement verifies the fluid inclusion Rb–Sr age. A noticeable feature of the isochron (Fig. 1) is the large range in $^{87}\text{Rb}/^{86}\text{Sr}$ which minimizes the influence of the initial $^{87}\text{Sr}/^{86}\text{Sr}$ on the calculated age. This is attributed to a variable input of Rb throughout the period of quartz deposition, either from a purely magmatic source or released into the ore fluids as a result of concomitant wallrock alteration (that is, greisenization). High Rb values are consistent with the preferential partition of Rb into magmatic–hydrothermal solutions¹³. Because the inclusions are largely pseudosecondary in origin, no two samples will contain identical proportions of the same generation of inclusion. Hence there will be a natural variation in the observed $^{87}\text{Rb}/^{86}\text{Sr}$ ratios.

The fluid inclusion approach seems superior to techniques that rely on the complete resetting of pre-existing systems (that is, the annealing of fission tracks in wallrock minerals adjacent to mineral veins¹⁴). Furthermore being relatively stable, inclusions are less sensitive to subsequent thermal events which often overprint the true age of mineralization. This is especially so for K–Ar studies on wallrock clays which show a poor retentivity for radiogenic argon. Of comparable significance is the ease with which K–Ar systematics can be reset by later meteoric–hydrothermal activity¹⁵. Because later hydrothermal solutions will often take advantage of earlier mineralizing channels, caution must be exercised when interpreting the K–Ar ages for feldspars and clays¹⁶. There is no *a priori* means of assessing whether a system has remained closed. It is also suspected that meteoric–hydrothermal activity may equally disturb Rb–Sr systematics. As yet the common Pb method for dating lead mineralization is only applicable to a small number of deposits of

the 'conformable' type. Recent refinements of terrestrial lead isotopic growth curves has done much to improve this technique and 'model' ages agree well with ages obtained by other methods. However, for most ore deposits, lead isotopic compositions indicate multistage histories which are not readily amenable to interpretation.

We thank R. Pankhurst for the development of programmes for the automatic running of the mass spectrometer, also D. Norman whose early isotopic work on fluid inclusions in Norwegian granites stimulated our research. The project was financed in part by the Commission of the European Communities Raw Materials Research Programme: Contract No. 038-79-1 MPP UK. Publication is by permission of the Director, Institute of Geological Sciences (NERC).

Received 23 December 1980; accepted 12 February 1981.

1. Roedder, E. & Skinner, B. J. *Econ. Geol.* **63**, 715–730 (1968).
2. Lokerman, A. A. *Mineralog. Sb. L'vov* **16**, 312–317 (1962).
3. Kerrich, R. *Contr. Miner. Petrol.* **59**, 195–202 (1976).
4. Skinner, B. J. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 1–21 (Wiley, New York, 1979).
5. Shepherd, T. J., Beckinsale, R. D., Rundle, C. C. & Durham, J. *Trans. Instn Min. Metall.* **85**, B63–B73 (1976).
6. McMahon Moore, A. J. *Jl R. Sch. Mines* **26**, 7–14 (1977).
7. Ermakov, N. P. in *Int. Ser. Monogr. Earth Sci.* **22**, 30–41 (1965).
8. Roedder, E. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 684–731 (Wiley, New York, 1979).
9. Faure, G. in *Principles of Isotope Geology* (ed. Faure, G.) 68–70 (Wiley, New York, 1977).
10. Williamson, J. H. *Can. J. Phys.* **46**, 1845–1847 (1968).
11. Wadge, A. J., Gale, N. H., Beckinsale, R. D. & Rundle, C. C. *Proc. Yorks. geol. Soc.* **42**, 297–305 (1978).
12. Brown, P. E., Miller, J. A. & Soper, N. J. *Proc. Yorks. geol. Soc.* **34**, 331–342 (1964).
13. Plimer, I. R. & Elliot, S. M. *J. Geochem. Explor.* **12**, 21–34 (1979).
14. Ashley, R. P. & Silberman, M. L. *Econ. Geol.* **71**, 904–924 (1976).
15. Taylor, H. P. Jr *J. geol. Soc. Lond.* **133**, 509–558 (1977).
16. Halliday, A. N. *Econ. Geol.* **75**, 752–759 (1980).

Biogeochemical significance of a novel sedimentary C_{27} stanol

S. C. Brassell & G. Eglinton

Organic Geochemistry Unit, University of Bristol,
School of Chemistry, Bristol BS8 1TS, UK

Recent studies^{1–3} of the lipids of marine organisms have revealed numerous novel sterols, many with modified side chains. As yet few of these compounds or their diagenetic products have been recognized in geological materials. However, recent investigations of various marine sediments^{4–8} have detected, for example, gorgosterol⁴ (IIo) and several 4-methylstanols^{7,8} (such as III/). Our gas chromatography (GC) and computerized gas chromatography–mass spectrometry (C-GC-MS) investigations of the extractable lipids in Neogene deep-sea sediments from the Japan Trench have revealed highly complex sterol distributions within which 69 different sterols have been characterized⁹. Several of these sterols^{9,10} (such as II_m, II_n and III_o), have been reported as constituents of marine organisms^{11–13}, but not of marine sediments; others are of unknown structure. Here we report the identification of one of these novel sterols, 27-nor-24-methyl-5 α -cholestan-3 β -ol (Ih) which supports the biological demethylation pathways proposed by Kobayashi and Mitsunashi¹². Its diagenetic products would also include 27-nor-24-methylcholestanes which may be unrecognized components of the sterane mass fragmentograms widely used by petroleum geochemists for oil and source rock fingerprinting. Such components might also prove valuable marine markers in the interpretation of palaeoenvironments.

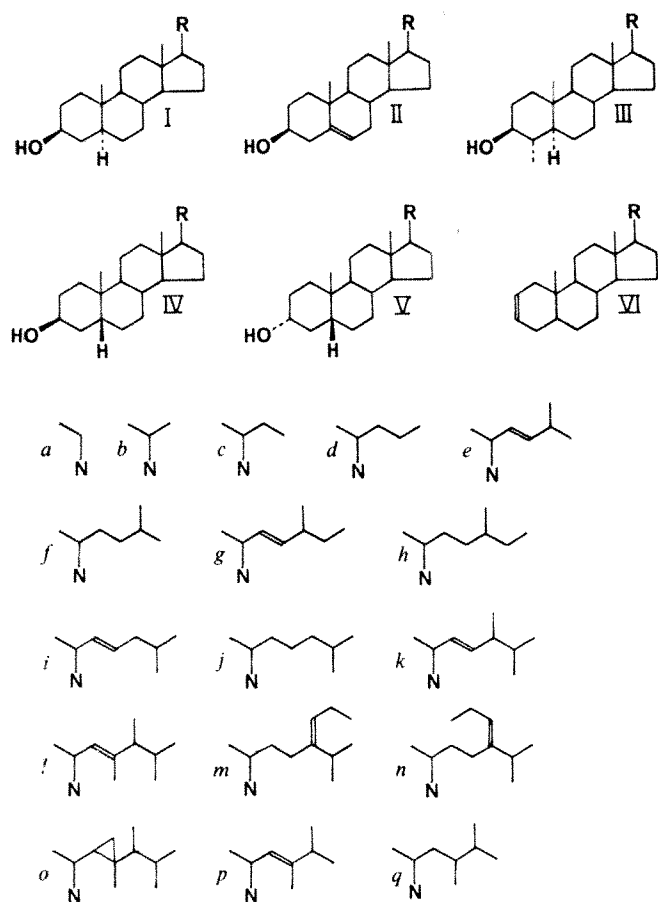


Fig. 1 Structural formulae of steroidal nuclei (N, I–VI) and side chains (R, a–q). The sterols I, II and III, bearing a wide range of side chains, are produced by organisms^{1–3,11–13,22–24} and occur in sediments^{4–10}. In contrast, the 5β -stanols (IV and V) and the Δ^2 -sterene (VI) are apparently derived from microbial and/or geochemical processes, and are often prominent components of immature sediments^{8,9,21,27,28}; however, the known variety of their side chains is less extensive than that of stenols and 5α -stanols.

During Legs 56 and 57 of the Deep Sea Drilling Project (DSDP), the Glomar Challenger drilled at seven sites in a transect of the Japan Trench^{14,15}. Our studies concentrated on the extractable lipids of 10 samples of Upper Pleistocene (~0.5 Myr) to Upper Miocene (~6 Myr) age, from four of these sites, with the aim of understanding the origin and diagenetic fate of the lipids^{9,10,16}. The analytical procedure used to obtain free sterol fractions from these sediments followed that previously described¹⁷, with subsequent modifications^{16,18}. Briefly, the extractable lipids obtained by sediment ultrasonication with solvent were divided into neutral and acidic components following saponification, and thin layer chromatography was used to separate the total neutrals further into fractions principally composed of hydrocarbon, ketone and alcohol constituents. Each fraction was analysed by capillary GC (SE-52 or OV-101 columns, deactivated by the barium carbonate technique of Grob *et al.*¹⁹) and C-GC-MS after derivatization (for example, of alcohols to their trimethylsilyl (TMS) ethers) when appropriate.

27-Nor-24-methyl- 5α -cholestan- 3β -ol (Ih) was identified from its GC retention time and mass spectral characteristics and by GC and C-GC-MS co-injection with a standard prepared by hydrogenation (PtO_2/H_2 in EtOAc) of ocellasterol (27-nor-24-methylcholest-5,22E-dien- 3β -ol, IIg). Its mass spectrum (as a TMS ether) is indistinguishable from that of 5α -cholestan- 3β -ol (Ij, Fig. 2), but the retention times of these two compounds differ (Fig. 3). Previous analyses of sedimentary sterols may not have

recognized 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) because of the prominence of 5α -cholestan- 3β -ol (Ij) in such distributions and the necessity of fast scanning GC-MS (for example, 1 s scans) to distinguish the two compounds satisfactorily.

In the Japan Trench sediments, which are principally diatomaceous oozes and clays, the concentration of 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih), with the exception of one sample that contained only trace amounts (<0.1 ng per g dry sediment), varied from 0.6 to 11.7 ng per g dry sediment; values that correspond to 0.2 to 1.2 μg per g organic carbon. In four samples 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) was more abundant than 5α -cholestan- 3β -ol (Ij), attesting its importance as a sedimentary sterol.

Although the origin of 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) is uncertain, it seems to be biological rather than geochemical. No apparent correlation exists between its abundance and that of the other 27-nor-24-methylcholesteroids with $\Delta^{5,22}$ (occellasterol (IIg)) and Δ^{22} (Ig) unsaturation, data that argue against its formation by geochemical reduction²¹. In addition, 27-nor-24-methylcholest-5-en- 3β -ol (IIh) would be an expected product of such a nonspecific reduction process, but it was not detected in any of the sediments. Furthermore, the presence of only two 5β -stanols, 5β -cholestan- 3β -ol (IVj) and 5β -cholestan- 3α -ol (Vj), rather than a range commensurate with that of the Δ^5 -stenols and 5α -stanols indicates that reduction of Δ^5 -stenols is not a significant process in Japan Trench sediments and suggests that the 5α -stanols in the C_{21} – C_{31} range⁹, are derived directly from biological inputs.

The identification of 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) means that both 27-nor-24-methylcholesteroids and 24-norcholesteroids are now known to occur as $\Delta^{5,22}$ -stenols and Δ^{22} - and 5α -stanols, supporting the demethylation pathways proposed by Kobayashi and Mitsuhashi¹². The organisms that act as the biological source of 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) are unclear. Its C_{26} homologue, 24-nor- 5α -cholestan- 3β -ol (If) has only been isolated from a sponge²²; both this compound and other sponge constituents, such as C_{21} – C_{24} 5α -stanols (Ia–d) (refs 23, 24), have been recognized in the Japan Trench sediments^{9,16}, where sponge spicules are also found^{25,26}. Hence, 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) may be of poriferan origin.

Considering the geological fate of 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih), at an early stage of diagenesis stanols can undergo dehydration, generating Δ^2 -sterenes^{27–29}. The Δ^2 C_{27} sterene in the Japan Trench sediments which eluted just after cholest-2-ene (VIj) and possessed an identical mass spectrum is, therefore, probably 27-nor-24-methylcholest-2-ene (VIh); if so, then the diagenetic reactions of 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) parallel those of 5α -cholestan- 3β -ol (Ij) (refs 8, 9, 27–29). Hence, 27-nor-24-methyl- 5α -cholestan- 3β -ol (Ih) could be a precursor for other sedimentary steroidal hydrocarbons with 27-nor-24-methyl side chains, particularly diasterenes³⁰, aromatic steroids³¹ and steranes, which might prove valuable indicators of marine sources of sedimentary organic matter. The possible existence of such 27-nor-24-methylcholestanes has a bearing on petroleum geochemical studies in view of the widespread use of steranes as biological markers³³. In particular, the process of diagenesis and maturation effect sterane isomerization, generating a wide range of stereoisomers^{32,33}; hence, mature sediments and oils would be expected to contain a comparable variety of 27-nor-24-methylcholestane isomers. Significant amounts of such compounds would markedly complicate the mass fragmentograms used for sterane evaluation^{32,33}, and also affect the quantitation of individual stereoisomers.

Novel compounds are likely to continue to be found in marine sediments because of the role of sediments as 'sinks' for organic matter. For example, the novel C_{27} $\Delta^{5,22}$ -stenol, Δ^{22} -stanol and 5α -stanol in the Japan Trench sediments may possess 26,27-bisnor-23,24-dimethyl side chains (IIp, Ip and Iq in Fig. 3), as recently suggested³ for similar components in Walvis Bay diatomaceous ooze⁴.

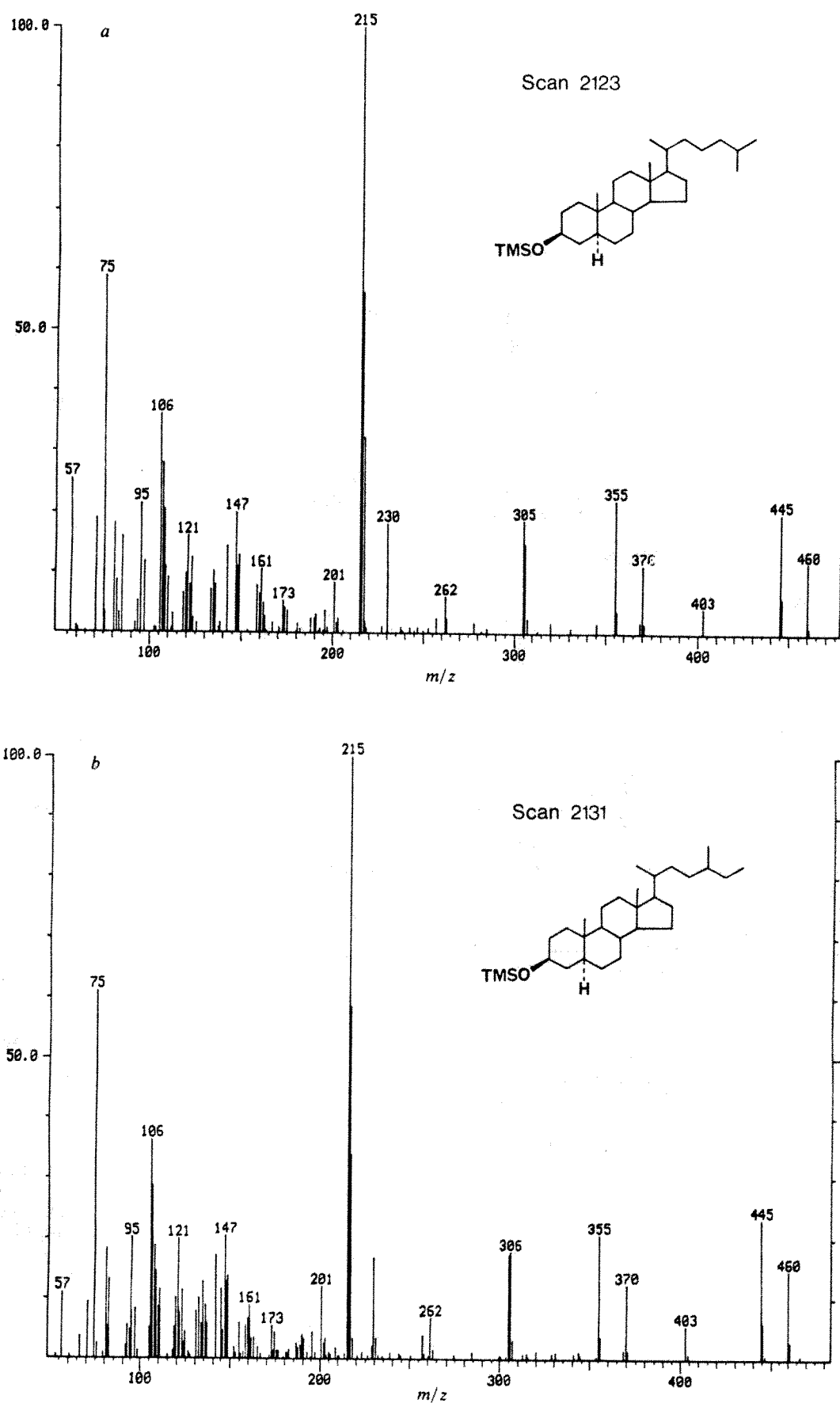


Fig. 2 EI mass spectra (40 eV, background subtracted) of: (a) 5α-cholestan-3β-ol (Ij); (b) 27-nor-24-methyl-5α-cholestan-3β-ol (Ih), as TMS ethers. The data were obtained from C-GC-MS analysis of the alcohol fraction separated from a Lower Pliocene diatomaceous claystone (DSDP Section 57-440B-39-3, 504m sub-bottom depth). The C-5 stereochemistry of (b) is assigned on the basis of its chromatographic characteristics and mass spectral fragmentation which are indicative of a 5α-stanol rather than a 5β-stanol⁸.

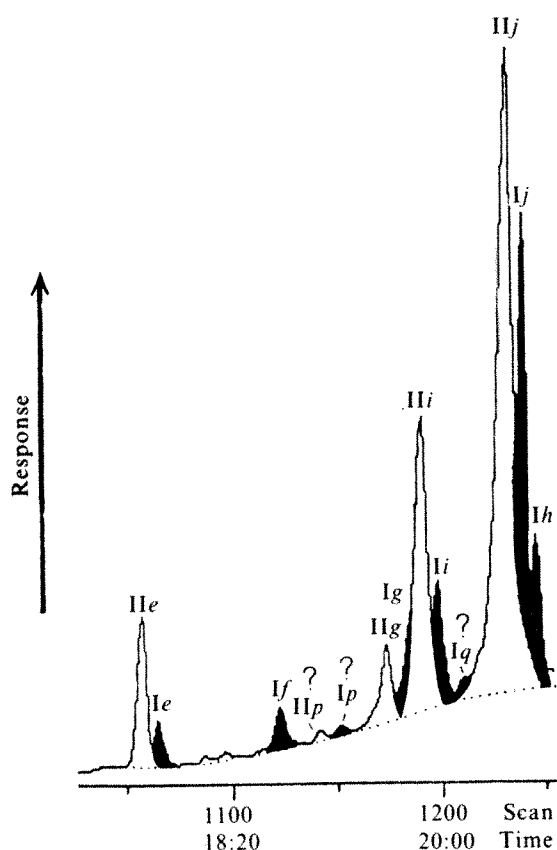


Fig. 3 Partial reconstructed ion chromatogram (RIC) from C-GC-MS analysis of extractable 4-desmethylsterols (as TMS ethers) in a Lower Pleistocene diatomaceous clay. □, Δ^5 -stenols; ■, 5α -stanols. The C_{26} and C_{27} sterol region is illustrated with designated peaks referring to the structures of Fig. 1. The 5β -stanols (IVj and Vj) found in this sample were present in the 4-methylsterol fraction (not shown). C-GC-MS conditions; 20 m OV-1 WCOT capillary column programmed, after splitless injection, from 50 to 100 °C at 6 °C min⁻¹, then 100 to 260 °C at 4 °C min⁻¹ and held isothermally at 260 °C. The mass spectrometer (Finnigan 4000 with INCOS 2300 data system) was scanned cyclically at 35 eV with complete mass spectra (m/z 50–560) being recorded every second. Data acquisition was initiated at 165 °C. A more extensive RIC of this sterol fraction is published elsewhere¹⁰.

We thank NERC (GR3/2951 and GR3/3758) for C-GC-MS facilities and a studentship (S.C.B.) and Dr M. Kobayashi for an authentic sample of ocellasterol. We acknowledge our colleagues associated with the DSDP project at Bristol and Mrs A. P. Gowar for assistance with C-GC-MS analyses.

Received 4 December 1980; accepted 10 February 1981.

- Minale, L. & Sodano, G. in *Marine Natural Products Chemistry* (eds Faulkner, D. J. & Fenical, W. H.) 87–109 (Plenum, New York, 1977).
- Schmitz, F. J. in *Marine Natural Products* Vol. 1 (ed. Scheuer, P. J.) 241–297 (Academic, London, 1978).
- Djerassi, C., Theobald, N., Kokke, W. C. M. C., Pak, C. S. & Carlson, R. M. K. *Pure appl. Chem.* **51**, 1815–1828 (1979).
- Wardroper, A. M. K., Maxwell, J. R. & Morris, R. J. *Steroids* **32**, 203–221 (1978).
- Lee, C., Gagosian, R. B. & Farrington, J. W. *Geochim. cosmochim. Acta* **41**, 985–992 (1977).
- Lee, C., Farrington, J. W. & Gagosian, R. B. *Geochim. cosmochim. Acta* **43**, 35–46 (1979).
- Boon, J. J. *et al. Nature* **277**, 125–127 (1979).
- Wardroper, A. M. K. thesis, Univ. Bristol (1979).
- Brassell, S. C. thesis, Univ. Bristol (1980).
- Brassell, S. C. *et al. in Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 375–392 (Pergamon, Oxford, 1980).
- Idler, D. R., Safe, L. M. & MacDonald, E. F. *Steroids* **18**, 545–553 (1971).
- Kobayashi, M. & Mitsuhashi, M. *Steroids* **26**, 605–624 (1975).
- Withers, N. W., Kokke, W. C. M. C., Rohmer, M., Fenical, W. H. & Djerassi, C. *Tetrahedron Lett.* 3601–3604 (1979).
- Langseth, M. *et al. Geotimes* **23**(3), 22–26 (1978).
- Von Huene, R. *et al. Geotimes* **23**(4), 16–21 (1978).
- Brassell, S. C. *et al. in Init. Rep. DSDP Leg 56/57*, 1367–1390 (1980).

- Barnes, P. J. *et al. in Init. Rep. DSDP Leg 48*, 965–976 (1979).
- Brassell, S. C. *et al. in Init. Rep. DSDP Leg 50*, 647–664 (1980).
- Grob, K., Grob, G. & Grob, K. Jr *Chromatographia* **10**, 181–187 (1977).
- Brassell, S. C., Gowar, A. P. & Eglinton, G. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 421–426 (Pergamon, Oxford, 1980).
- Gaskell, S. J. & Eglinton, G. *Nature* **254**, 209–211 (1975).
- Delseth, C., Tolela, L., Scheuer, P. J., Wells, R. J. & Djerassi, C. *Helv. chim. Acta* **62**, 101–109 (1979).
- Delseth, C., Carlson, R. M. K., Djerassi, C., Erdman, T. R. & Scheuer, P. J. *Helv. chim. Acta* **61**, 1470–1476 (1978).
- Carlson, R. M. K. *et al. Bioorg. Chem.* **7**, 453–479 (1978).
- Langseth, M. *et al. Init. Rep. DSDP Leg 56/57 Part 1* (1980).
- Von Huene, R. *et al. Init. Rep. DSDP Leg 56/57 Part 1* (1980).
- Dastillung, M. & Albrecht, P. *Nature* **269**, 678–679 (1977).
- Gagosian, R. B. & Farrington, J. W. *Geochim. cosmochim. Acta* **42**, 1091–1101 (1978).
- Gagosian, R. B., Smith, S. O., Lee, C., Farrington, J. W. & Frew, N. M. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 407–419 (Pergamon, Oxford, 1980).
- Rubinstein, I., Sieskind, O. & Albrecht, P. *JCS Perkin I* 1833–1835 (1975).
- Schaeffle, J., Ludwig, B., Albrecht, P. & Ourisson, G. *Tetrahedron Lett.* 4163–4166 (1978).
- Mackenzie, A. S., Patience, R. L., Maxwell, J. R., Vandenbroucke, M. & Durand, B. *Geochim. cosmochim. Acta* **44**, 1709–1721 (1980).
- Seifert, W. K. & Moldovan, J. M. *Geochim. cosmochim. Acta* **42**, 77–92 (1978) & **43**, 222–236 (1979).

Rate of dune reddening in a humid tropical climate

K. Pye

Department of Earth Sciences, University of Cambridge,
Downing Street, Cambridge CB2 3EQ, UK

The formation of red beds has long interested geologists and pedologists^{1–3}. It has been suggested that the nature and degree of reddening in dune sands can be a useful indicator of the relative age and environmental history of the sands; several authors have concluded that dune sands undergo progressive reddening with time^{4–7}, and James and Stanley⁸ considered that coastal dunes might actually be mapped on the basis of colour. However, quantitative data regarding the rate of reddening in different environmental conditions are generally lacking. I present here radiocarbon evidence from north-east Australia which suggests that red weathering profiles have developed in stabilized dunes of Holocene age in humid tropical conditions within the past 7,500 yr.

Deeply weathered siliceous sand dunes occur on several parts of the North Queensland coast, Australia, but are most extensive around Cape Bedford, Cape Flattery, and Shelburne Bay (Fig. 1). The dunefields consist of several generations of elongate parabolic dunes, the youngest of which are still partly active. The older stabilized dunes are deeply podsolized, with typical profiles on dune crests consisting of 1–4 m of leached white sand overlying 6 m or more of red and orange ferruginous sand (Fig. 2). The colour of the sands below the A horizons is variable, with banding and mottling characteristic of the lower profile, but the reddest B horizons attain 2.5YR4/8 (moist) on the Munsell soil colour chart. The red colour is due to iron oxides and hydroxides which cannot be positively identified by X-ray diffraction but which probably consist largely of poorly crystallized forms of haematite containing a certain amount of water. Bulk samples of red sand contain up to 0.668% of dithionite citrate extractable Fe, together with up to 0.75% extractable Al. The pigment is present both as coatings on framework grains and as interstitial matrix.

Air photograph interpretation and field drilling have shown that at least three distinct episodes of dune formation are represented at Cape Flattery. Dunes belonging to both the first and second episodes show evidence of reddening, but the colouration is most marked in dunes of the second generation. The forms of the oldest dunes are highly degraded and the sands contain black horizons of organic matter-cemented sandrock which are particularly well developed in low-lying depressions.

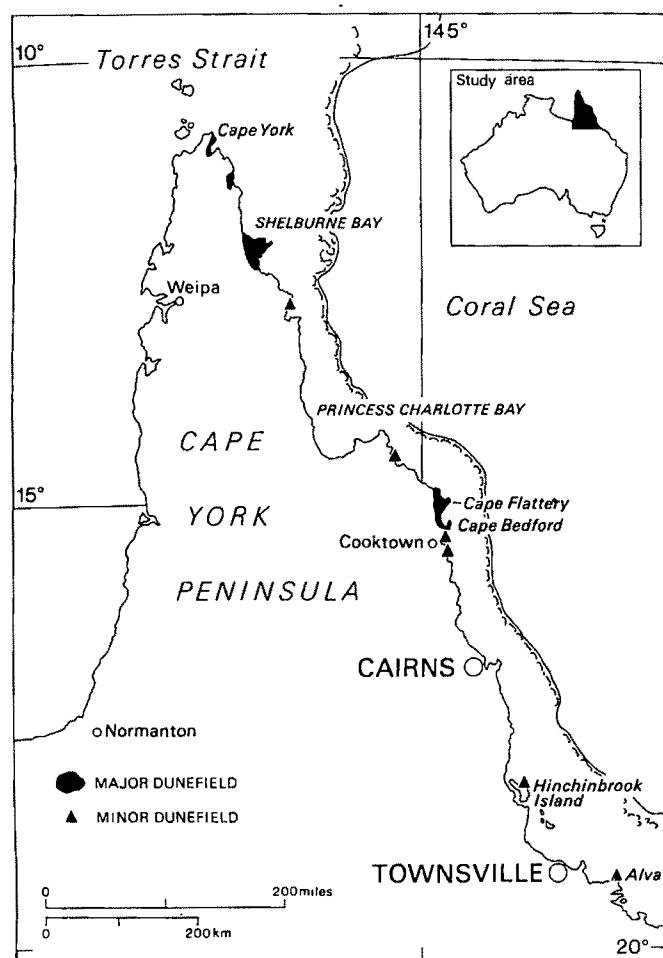


Fig. 1 Map showing location of Cape Flattery and other coastal dunefields in North-east Australia.

Some dunes show that part of the red pigment is being removed by progressive podsolization. A radiocarbon date obtained from a large root in growth position and embedded in humate⁹ exposed in coastal cliffs 3 km north-west of Cape Flattery indicated a 'background' age of 48,000 ^{14}C yr (Q-2085). Similar deposits occur in Southern Queensland¹⁰ and New South Wales where they have been shown to be last interglacial in age¹¹.

The forms of the second generation of dunes, which drilling has shown to overlie partly the first generation, are relatively sharp and the sands contain only very thin, non-indurated humate horizons. A typical weathering profile found in dunes of this generation is shown in Fig. 2. Two radiocarbon dates have been obtained from samples of charcoal collected at depths of 2 and 4 m respectively within this profile. The charcoal was pretreated with an alkali-pyrophosphate mixture followed by acidification to remove humic acid contaminants. The upper charcoal sample gave a date of $7,480 \pm 75$ ^{14}C yr BP (Q-2081), and the lower sample indicated an age of $8,200 \pm 85$ ^{14}C yr BP (Q-2082). As the charcoal fragments must have been incorporated within the sands while the dune was active, the dates suggest that stabilization and weathering has taken place within the past 7,500 ^{14}C yr. The possibility that the charcoal is reworked from older deposits cannot be entirely ruled out, and in this event the dates would indicate a maximum age for the reddened dune.

The youngest dune sand unit at Cape Flattery includes all the active and recently active dunes which have partly buried and reworked the older sands. These dunes mostly consist of undifferentiated white quartz sand, but those which have been stabilized longest show immature podsol development with brown or brownish-orange B horizons. Buried organic Al

horizons are exposed in the walls of some dunes. Two radiocarbon dates on *in situ* wood stumps from one such buried soil including a 1 m-thick brownish B horizon both indicated 'modern' ages (SRR-1534 and SRR-1535).

Previous work on red sediments has led to the conclusion that reddening is generally slow, and in arid regions may take 10^5 – 10^6 yr (refs 12, 13). However, the evidence reported here suggests that in humid tropical conditions bright red colours may be attained within 10^3 – 10^4 yr. Furthermore, over a longer time period much of the initial red colour may be lost as humic and fulvic acids accumulate within the profile. A similar conclusion was reached, though without radiocarbon evidence, by Andriess¹⁴ in relation to reddened beach ridges and alluvial deposits in Sarawak.

Rapid reddening of stabilized dunes at Cape Flattery is favoured by several factors. First, abundant detrital ilmenite and magnetite grains within the sands provide a large potential source of iron which can be released by weathering^{15–19}. Second, rapid release of iron and formation of ferric oxides is favoured by oxidizing interstitial conditions which reflect the good drainage of the dune sands and the seasonally humid tropical climate. At Cooktown, located just south of Cape Flattery, mean daily temperature is above 25 °C throughout the year, reaching 31.3 °C in December, while about 75% of the mean annual rainfall occurs between January and April. Third, weathering of ilmenite, magnetite and other minerals may be enhanced by organic acids released by the heath and woodland vegetation characteristic of the dune areas^{20,21}. Soil pH values at Cape Flattery generally lie in the range 3.5–5.5.

Rate of reddening in dune sands and other recent sediments is dependent on several variables including sand mineralogy, temperature, moisture availability, drainage conditions, degree of sediment stability and significance of airborne dust input²². Climatic conditions undoubtedly permit reddening in both the arid and humid tropics, but, other things being equal, the time required to attain a given degree of redness is likely to be longer in the arid tropics due to lower moisture availability, a slower rate of mineral weathering and relative unimportance of biochemical processes. In humid tropical areas periods of seasonal

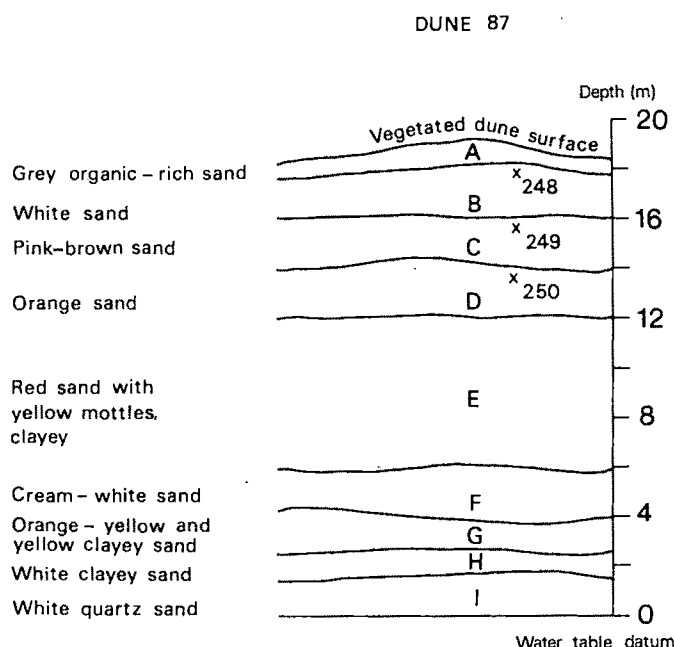


Fig. 2 Diagrammatic representation of a weathering profile exposed in a section through a stabilized dune near Cape Flattery. Radiocarbon dates Q-2081 and Q-2082 refer to charcoal samples (x) 248 and 249 respectively. Charcoal sample 250 was analysed for chemical purposes.

drought and water table lowering can also be expected to favour the formation of haematite²³. However, owing to the diverse nature of the controls on reddening, time cannot be regarded as an independent variable in the reddening process, and colour alone cannot be used as an adequate guide to age or morpho-stratigraphic correlation between different areas.

This research was supported by NERC and the RS. Radiocarbon age determinations were made by Dr Roy Switsur and by the NERC Radiocarbon Dating Laboratory, East Kilbride.

Received 9 January; accepted 6 February 1981.

1. Krynine, P. D. *Trans. N.Y. Acad. Sci.* **11**, 60–68 (1949).
2. Van Houten, F. B. in *Problems in Palaeoclimatology* (ed. Nairn, A. E. M.) 647–661 (Interscience, New York, 1964).
3. Turner, P. *Continental Red Beds* (Elsevier, Amsterdam, 1980).
4. Price, W. A. *Bull. geol. Soc. Am.* **73**, 1281–1289 (1962).
5. Norris, R. M. *J. sedim. Petrol.* **39**, 7–11 (1969).
6. Dolan, R. J. *J. sedim. Petrol.* **40**, 765 (1970).
7. Walker, T. R. *US geol. Surv. Prof. Pap.* 1052, 61–81 (1979).
8. James, N. P. & Stanley, D. J. *Smithson. misc. Collns* **152**, 1–31 (1967).
9. Swanson, V. E. & Palacas, J. G. *US geol. Surv. Bull.* 1214-B (1965).
10. Ward, W. T., Little, I. P. & Thompson, C. H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **26**, 305–316 (1979).
11. Marshall, J. F. & Thom, B. G. *Nature* **263**, 120–121 (1976).
12. Walker, T. R. *Bull. geol. Soc. Am.* **78**, 353–368 & 917–920 (1967).
13. Walker, T. R., Ribbe, P. H. & Honea, R. M. *Bull. geol. Soc. Am.* **78**, 1055–1060 (1967).
14. Andriess, J. P. *Geoderma* **3**, 261–279 (1969–70).
15. Temple, A. K. *Econ. Geol.* **61**, 341–348 (1966).
16. Reynders, J. J. *Nova Guinea* **6**, 1–159 (1964).
17. Colombo, U., Fagherazzi, G., Gazzanini, F., Lanzavecchia, G. & Sironi, G. *Nature* **202**, 175–176 (1964).
18. Gallagher, K. J., Feitrecht, W. & Mannweiler, U. *Nature* **217**, 1118–1121 (1968).
19. Van Houten, F. B. *Bull. geol. Soc. Am.* **83**, 2761–2772 (1972).
20. Pye, K. & Jackes, B. *Proc. R. Soc. Qd* **92**, 37–42 (1981).
21. Schnitzer, M. & Khan, S. U. *Humic Substances in the Environment* (Dekker, New York, 1972).
22. Gardner, R. M. & Pye, K. *Prog. phys. Geogr.* (in the press).
23. Schmalz, R. F. *Bull. geol. Soc. Am.* **79**, 277–280 (1968).

Internal plumbing of Mount Etna

J. E. Guest

University of London Observatory, Mill Hill Park,
London NW7 2QS, UK

A. M. Duncan

Geology Section, Luton College of Higher Education,
Luton LU1 3JU, UK

A number of conflicting models have been proposed for the internal plumbing system of Etna; some of the conflict arises from attempts to fit the same model to the whole of Etna's history. We argue that at present magma reaches the surface without significant high-level storage during its ascent, and that eruptions are tectonically controlled. These conditions have probably prevailed for the past 200 yr. However, evidence from lavas erupted in prehistoric times suggests that substantial low-pressure differentiation accompanied by different styles of eruption took place, implying a different plumbing system which includes high-level storage of magma.

Etna is a central-vent volcano some 3 km high, with numerous parasitic cones. The position of the central conduit has shifted with time, each change in location generally being preceded by caldera collapse¹. Almost continuous activity occurs at or near the summit where there is a steady effusion of lava, normally at the rate of less than 1 m³ s⁻¹, accompanied by strombolian activity. A curious aspect of summit activity is that two craters less than a few hundred metres apart may simultaneously show quite different activity; for example, in 1966 to 1971 the north-

east crater was erupting lava on the surface, while only 200 m away, activity in the chasm was ~1,000 m below this level. Persistent summit activity usually stops when flank eruptions occur—these are characterized by much higher rates of effusion, typically at tens of cubic metres per second. Feeding fissures tend to parallel the complex regional tectonic pattern.

Based on petrological evidence, Rittmann² envisaged that magma is supplied under Etna by abyssal fissures and that there is no magma chamber. Guest³ suggested the existence of a plexus of dykes and sills within or just below the volcanic pile, analogous to that proposed for Kilauea⁴. From horizontal ground deformation studies between 1971 and 1974, Wadge^{5,6} postulated two magma reservoirs: one at the base of the volcanic pile, and the other an open cylindrical reservoir in the upper part of the central conduit. The central conduit reservoir, known as the chasm, was envisaged by Wadge as feeding flank eruptions by first opening fissures radial to the chasm as a result of hydraulic fracturing and subsequent drainage of the reservoir under gravity through these fissures. Tanguy and Kieffer⁷ suggested high-level magma storage in active dyke-like intrusions, eruption from which was triggered by tectonic activity.

Studies of vertical ground deformation have been made between 1975 and 1980⁸. During this time persistent activity occurred from the north-east crater and the north-east rift between 1975 and 1978, and flank eruptions occurred on the southern flank between May 1978 and August 1979. Following the established deformation pattern at Kilauea, while relatively little deformation might have been expected from some of the persistent activity at the summit, substantial effusion low on the volcano flanks should have been preceded by inflation caused by high-level magma storage, followed by deflation as the chamber emptied. However, during 1975–80, including the major August 1979 flank eruption during which an estimated minimum of 8 × 10⁶ m³ of magma was erupted (C. Kilburn, personal communication), no gross pattern of inflation or deflation was observed either by precise levelling traverse across the summit, or from some 35 dry-tilt stations. Based on five years of tilt measurement on Etna, there is no evidence that large amounts of magma are stored within the volcano before eruption, and if it is stored in an intermediary reservoir, this lies well below the base of the volcano. Thus internal conditions at Etna seem to differ from those of Kilauea.

Although no gross accumulation of magma within the Etna volcano can be identified, small vertical movements do occur. A 1.5-cm inflation area of ~1 km radius was observed two years before the May 1978 eruption in the region of the main vent, suggesting that a narrow wedge of magma was emplaced below the eruption site before effusion started⁹.

Although the vertical deformation results do not support a high-level magma chamber in a conventional sense, they are not inconsistent with Wadge's⁶ proposal of an open chamber in the chasm at the top of the central conduit. However, Wadge's model requires that the chasm drains during the course of flank activity and that the amount of lava drained corresponds to the volume of lava erupted. These conditions have not been fulfilled by recent eruptions when, for example in April 1971, the chasm floor was lower than the active flank vents³. Thus, although the Wadge model is not discounted as a potential mechanism, it cannot be considered as a general model to explain Etna's plumbing system during 1970–80 (see also ref. 9).

Based on seismic evidence, Sharp *et al.*¹⁰ suggested a magma storage region at a depth of ~20 km below Etna. This has the form of a prolate ellipsoid with a width ~20–30 km and the long axis paralleling the north-east fissure trend through the mountain. From the same experiment they identified the Moho at 27 km, magma generation occurring below this level.

During historical times the lavas have been basic hawaiites with little variation in chemical composition^{11–13}. Historic lavas are strongly porphyritic with 30–40% phenocrysts of plagioclase, augite and olivine, plagioclase being the dominant phenocryst phase; the magma therefore undergoes extensive

crystallization before eruption. Limited variation in chemistry between lavas indicates that there is little differentiation. The predominance of plagioclase suggests that most of the crystallization takes place at moderate to low pressures. Experimental work¹⁴ on an alkali basalt shows that crystallization of plagioclase is suppressed to near solidus temperatures at pressures approaching 9 kbar, equivalent to ~22 km depth. These observations imply that considerable crystallization is taking place at relatively high levels, but the magma is erupted before significant differentiation occurs. This is consistent with the ground deformation interpretation that there is no substantial high-level magma storage. Magma stored in the open central conduit may, however, undergo vapour-phase fractionation¹⁵⁻¹⁸.

Although the composition of historical lavas has remained almost constant, there are petrographic distinctions, for example the 1974 flank eruption lavas^{7,19,20}, where plagioclase was not a primary phenocryst phase. In this case either the crystallization of plagioclase was suppressed by high P_{H_2O} conditions, or magma migrated from depth rapidly containing only early formed phenocrysts of augite and olivine.

The evidence given above supports the view expressed by Rittmann² that on Etna magmas ascend directly to the surface

from depth without storage in a high-level magma chamber. These observations agree to some extent with the model proposed by Tanguy and Kieffer⁷ of a system of active intrusions. It is likely that magma enters the volcanic pile permissively and a high level of magma in an open central conduit gives steady persistent activity. Dyke-like intrusions leading to flank eruption radiate at various depths from the central conduit and may be emplaced in tectonically dilated fissures. Eruptions are controlled both in space and time by the interaction of magma fluid pressure and the regional stress pattern (see ref. 21): a decrease in the minimum horizontal principal stress will increase the gape of dilational fissures, allowing an increased rate of magma uprise. The presence or absence of a cinder cone over the vent, for flank eruptions, will depend on the depth of connection between the feeder dyke and the conduit (Fig. 1a). Eccentric cones result from ascent of magma independent of the central conduit, for example, the Moio cone 18 km north of the summit²². The role of the 20-km deep magma reservoir is uncertain, but the observation that repose time between the end of a flank eruption and the onset of persistent activity is proportional to the volume of erupted material in the flank eruption⁶ may suggest that the repose period is controlled by the time taken for the reservoir to be replenished.

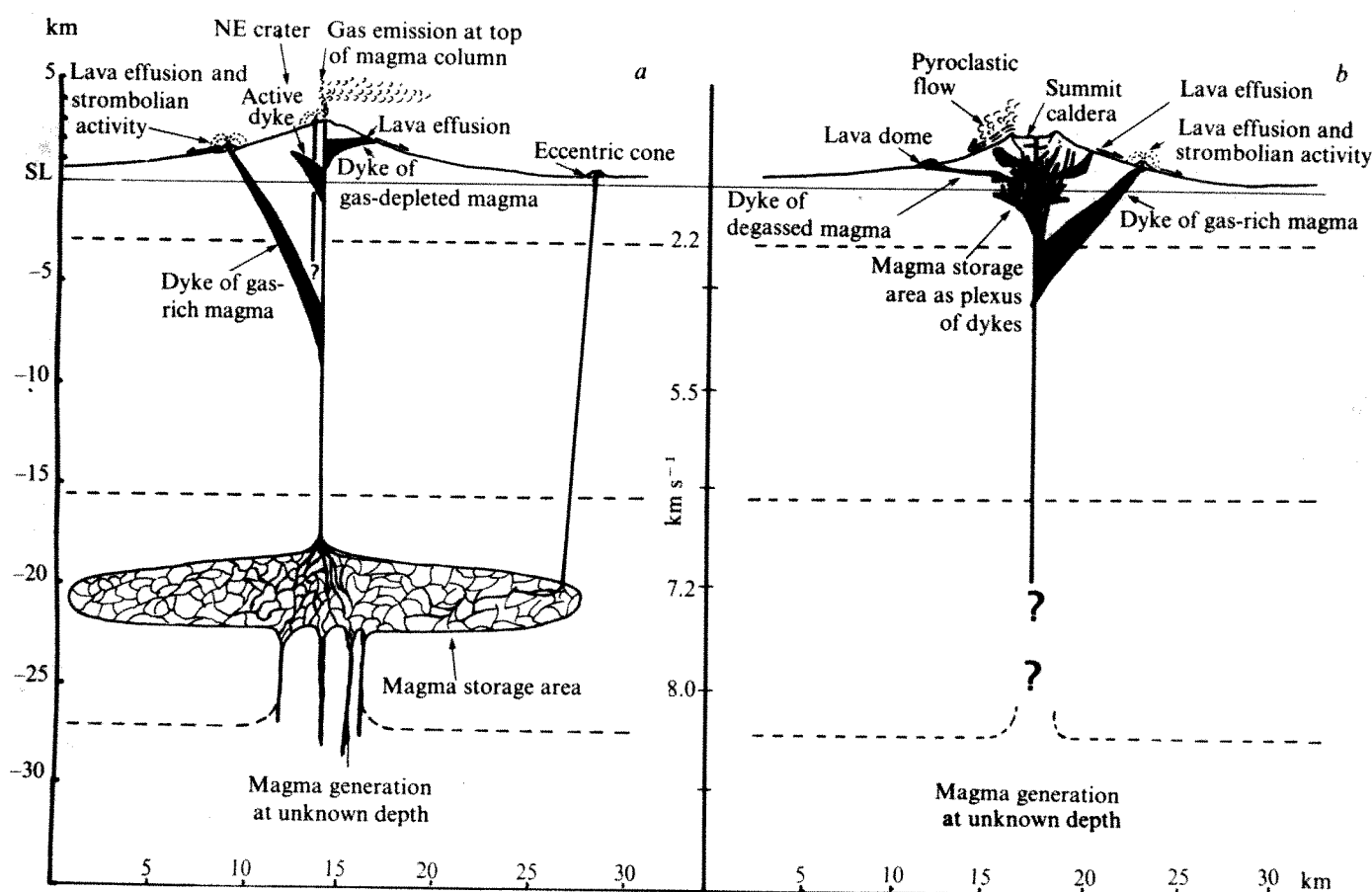


Fig. 1 a, Interpretation of the present internal structure of Mount Etna. Magma rises up the central conduit from a deep magma storage area. Low effusion rate persistent activity from this conduit occurs near the volcano summit. A column of magma is probably always present in the central conduit and associated conduit of the north-east crater, at levels of less than a kilometre below the volcano's summit. The opening of fissures by tectonic activity allows magma to leave the central conduit at different depths to produce flank eruptions with high effusion rates compared to persistent activity. Occasionally eccentric eruptions may occur as a result of magma rising directly from the deep magma storage area using a conduit not related to the central conduit. The dimensions of the magma storage area and the seismic velocity profile are from ref. 10. b, Condition of the volcano during pre-Valle del Bove times, showing high level plexus of dykes in which magma is stored. Although the different styles of activity are shown on the same diagram, the normal activity was probably strombolian with occasional extrusions of volcanic domes and paroxysmal events.

Geophysical information is only available for the past few years on Etna. Nevertheless, interpretations of internal plumbing based on these data are probably relevant for the past 200 yr, during which time the petrology of the lavas, styles of eruption and output of the volcano have been relatively constant. However, before 1669 the output of the volcano was considerably higher²² and the erupted lavas petrographically different from the more recent ones, although of hawaiitic composition. These earlier lavas are characterized by larger phenocrysts of plagioclase, typically 4 mm long, compared with those in later lavas which are ~1 mm long. The presence of fewer phenocrysts may imply a short period of high-level storage. A plexus of active dykes near the base of the volcanic pile may have existed at this time⁶. Thus, even during historic times the internal plumbing of Etna seems to have changed; these alterations may have resulted from variations in the regional stress pattern.

At times in the prehistoric past of Etna (5,000–100,000 yr BP), before formation of the Valle del Bove¹, a very different system of internal plumbing may have operated in the volcano. Among the older volcanics, evolved lavas of mugearitic and benmoreitic composition are relatively common, indicating more extensive differentiation of the ascending magma^{11–13}. Plagioclase is considered to have played a major part in this crystal fractionation¹⁸ suggesting that at one time there was high-level magma storage. Ash fall and ignimbrite deposits occur^{23,24}, indicating violent explosive events in the summit region, possibly associated with caldera collapse and de-gassing of a high-level magma reservoir (Fig. 1b). On the lower flanks of the volcano, relatively degassed magma was erupted, which formed benmoreitic autoclastic domes¹⁸ probably fed from a storage area that had been degassed by loss of volatiles up the central conduit.

We make the following conclusions: (1) there is no simple model of volcano plumbing that can be applied to central basaltic volcanoes, and a long-lived central volcano made up of repeated eruptions may not have the same type of plumbing throughout its history. (2) Petrogenetic models based on a whole suite of lavas spanning a long time period may be misleading as in this case, where different models involving the presence or absence of a high-level magma reservoir are applicable to different phases in the volcano's plumbing history. (3) Because a volcano's plumbing is controlled by factors such as the regional stress pattern, rate of magma supply and rate of stress adjustment in the magma source region, a study of Etna's changing internal conditions may have direct relevance to the regional tectonic evolution.

This work was funded by NERC and the Royal Society. We thank G. Wadge, C. Kilburn, J. B. Murray and T. Sanderson for helpful comments.

Received 27 November 1980; accepted 5 February 1981.

1. Guest, J. E. *U.K. Research on Mount Etna, 1977–1979* 7–8 (Royal Society, London, 1980).
2. Rittmann, A. *Phil. Trans. R. Soc. A274*, 5–16 (1973).
3. Guest, J. E. *Phil. Trans. R. Soc. A274*, 64–78 (1973).
4. Eaton, J. P. & Murata, K. *J. Science* **132**, 925–938 (1960).
5. Wadge, G. *J. Volcan. Geotherm. Res.* **1**, 237–264 (1976).
6. Wadge, G. *J. Volcan. Geotherm. Res.* **2**, 361–384 (1977).
7. Tanguy, J.-C. & Kieffer, G. *Bull. volcan.* **40–4**, 1–14 (1976/77).
8. Murray, J. B. & Guest, J. E. *U.K. Research on Mount Etna, 1977–1979* 18–20 (Royal Society, London, 1980).
9. Kasser, M., Delorme, H., Kieffer, G. & Tanguy, J.-C. *C.r. hebdom. Séanc. Acad. Sci., Paris* **287**, 1191–1194 (1978).
10. Sharp, A. D. L., Davis, P. M. & Gray, F. *Nature* **287**, 587–591 (1980).
11. Cristofolini, R. *Phil. Trans. R. Soc. A274*, 17–35 (1973).
12. Romano, R. & Sturiale, C. *Boll. geol. Soc. Ital.* **94**, 1109–1148 (1975).
13. Romano, R. & Guest, J. E. *Boll. geol. Soc. Ital.* **98**, 189–215 (1979).
14. Green, T. H., Green, D. H. & Ringwood, A. E. *Earth planet. Sci. Lett.* **2**, 41–51 (1967).
15. Romano, R. *Bull. volcan.* **34**, 823–832 (1970).
16. Romano, R. & Sturiale, C. *Phil. Trans. R. Soc. A274*, 37–43 (1973).
17. Rittmann, A. *Bull. volcan.* **38**, 1–17 (1974).
18. Duncan, A. M. *Geol. Mag.* **115**, 273–285 (1978).
19. Guest, J. E. *et al. Nature* **250**, 385–386 (1974).
20. Bottari, A., Lo Guidice, E., Patane, G., Romano, R. & Sturiale, C. *Riv. Mineraria Siciliana* **154–156**, 175–198 (1975).
21. Shaw, H. R. *Physics of Magnetic Processes* (ed. Hargraves, R. B.) 201–232 (Princeton University Press, 1980).
22. Wadge, G., Walker, G. P. L. & Guest, J. E. *Nature* **255**, 385–387 (1975).
23. Kieffer, G. *C.r. hebdom. Séanc. Acad. Sci., Paris* **278D**, 1549–1552 (1974).
24. Duncan, A. M. *Geol. Mag.* **113**, 357–363 (1976).

Anti-predator function of bluegill sunfish nesting colonies

Wallace J. Dominey

Section of Neurobiology and Behavior, Cornell University, Ithaca, New York 14853, USA

Although the adaptive significance of nesting in colonies has been investigated repeatedly in avian species, until now no detailed studies have been reported on colonial nesting in fishes. Investigators of avian colonial systems have suggested that birds nest in colonies either as a response to predators^{1,2} or to locate more easily ephemeral food resources^{3,4}. The avian literature is unclear, however, as to which selective pressure (predation or food resources) is more crucial, with investigators using similar evidence to reach essentially opposite conclusions⁵. I report here evidence obtained from a long-term study of a colonial fish, the bluegill sunfish (*Lepomis macrochirus*), showing extreme synchrony in breeding and the additional protection afforded offspring by tightly clustered groups of adults, which argues strongly that colonial nesting in this species and other fishes has evolved in response to predation pressure.

Observations on the reproductive system of the bluegill sunfish were made during the summers of 1976, 1978 and 1979 while skin or scuba diving in Lake Cazenovia, New York. In this lake, male bluegills construct nests (bowl-shaped depressions in the substrate) in colonies ranging in size from 15 to 500 nests. Throughout the breeding season males periodically reoccupy, breed in, and abandon these nesting areas. Although suitable nest sites are not limited (in preparation), during the spawning period occupied nests are packed rim to rim, with most nests having five to seven contiguous neighbours (Fig. 1).

As early as 1938, Darling⁶ suggested that colonially nesting birds might breed synchronously to reduce the time their young are available to local predators. However, in avian systems, an alternative hypothesis also predicts breeding synchrony. If nesting colonies serve as 'information centres' that enable unsuccessful or inexperienced foragers to follow more successful individuals to ephemeral food sources, individuals breeding in synchrony would gain an advantage⁷. Because colonially nesting fishes typically neither feed their young nor leave the colony to forage, the competing hypothesis of social foraging does not apply. Consequently, the presence of breeding synchrony in these organisms would provide stronger support for the predation hypothesis than synchronous breeding in avian systems.

If predation is the evolutionary force favouring colonial nesting in bluegills, individuals within a colony should synchronize their reproductive efforts, resulting in greater breeding synchrony within colonies than between colonies. As predicted by this hypothesis, breeding within bluegill colonies in Lake Cazenovia is remarkably synchronous. At a particular colony breeding episodes occur periodically, and consist of the establishment of males on nests, the arrival of gravid females which remain in the water column until spawning is initiated, and a brief 1- or 2-day spawning period. Between these spawning periods gravid females are absent and successful males provide parental care. Although common responsiveness to optimal breeding conditions seems to promote a degree of breeding synchrony among all the colonies in Lake Cazenovia, synchrony within colonies is much greater than between colonies, as predicted, and even neighbouring colonies can cycle out of phase.

The period immediately following the termination of parental care is often critical to an individual's survival. During this period, bluegill fry minimize their availability to predators by abandoning their nests synchronously, with most of the offspring in an entire colony leaving within a period of a few hours. Analogous fledgling synchrony has been reported in some

Fig. 1 An artist's conception of the activity in a bluegill colony during the spawning period. An aggregation of females is shown above a spawning centre (top centre). Multiple spawning partners (centre right) can indicate the presence of a 'female mimic'¹², a small sexually mature male that mimics female behaviour while 'stealing' fertilizations from nesting males. Nesting males are shown mobbing a large predatory turtle (top right), fanning eggs (right middle) and swimming rapidly around the perimeter of their nests in 'rim circling' displays (bottom right).



colonial birds⁸. In fish, as in birds, the impact of predation by locally available predators should be reduced when large numbers of offspring become available over a short period of time (as many as 100,000 in a single bluegill nest). The extreme synchrony observed both in breeding and nest abandonment argues strongly that colonial nesting in bluegills serves an anti-predator function.

In addition to synchronous breeding, active defence can limit the availability of the offspring of colonially nesting animals to predators. In a colony where each parent defends an area around its nest, if defended areas overlap, the offspring might benefit from the cumulative defence of several individuals. The offspring of colonially nesting birds have been found to gain such an advantage, with nests in high-density areas being better protected than those in low-density areas⁵, and this type of advantage has also been suggested for colonially nesting fishes⁹⁻¹¹. In the case of bluegills, evidence of a colony structure that promotes overlap of defended areas would support further the hypothesis that colonial nesting is a response to predation pressure, as would evidence that offspring in high-density areas gain additional protection.

Female bluegills select mates such that males guarding nests are found in more tightly packed groups than one would predict if females were indiscriminately selecting mates. During the spawning period, gravid females are attracted to spawning pairs and deposit their eggs in adjacent nests, creating areas of spawning activity or 'spawning centres'^{12,13}. Bluegill males whose nests are not located in these spawning centres often fail to receive eggs, or receive too few eggs to merit parental care, and abandon their nests following the spawning period. Might the spawning response of female bluegills function to ensure that their offspring are adequately protected from predation by a tightly compacted group of defending males?

Attempted predation was common in the bluegill colonies of Lake Cazenovia, occurring more frequently than has been reported in avian colonial systems. In a typical 10-min observation period, a bluegill male defended its nest several times against potential egg predators. In 950 min of cumulative observation, the most common piscivore egg predators were foraging (not nesting) conspecifics (72%) and members of the closely related species *Lepomis gibbosus*, the pumpkinseed sunfish (21%).

To test whether the offspring in nests in tightly clustered groups receive additional protection against predators, I

removed males from both isolated and surrounded nests, then recorded the number of egg predators present in the nests at 1-min intervals for 10 min. Nesting males were not observed to forage in temporarily vacated nests. Isolated nests (having no more than one neighbouring nest) were more vulnerable to egg predators than were surrounded nests (having five or more neighbouring nests) (Fig. 2). Not only did the presence of defending neighbours make it more difficult for foraging fishes to enter a focal nest, but neighbouring males often actively attacked egg predators foraging in the focal parentless nests, with isolated nests receiving significantly fewer defences by neighbours than surrounded nests (Fig. 2). The additional protection afforded to nests in high-density areas is expected to result both from the 'selfish herd' effect¹⁴, by which individuals passively gain from the presence of neighbours, and from the overlap of defended areas. Both protective effects should extend readily to other types of predators attacking bluegill offspring.

The biological significance of these findings is supported by two observations. (1) Offspring in parentless nests sometimes survive to the freeswimming stage because of the incidental protection of neighbouring males. (2) Foraging schools composed of bluegills and pumpkinseeds sometimes harass isolated nesting males until these males are no longer effective at nest defence; this results in complete loss of brood. I never observed the nests of males in high-density areas to be similarly overrun, apparently because of the incidental protection provided by neighbouring nest-defending males. The inability of isolated males to defend their nests against egg predators, especially those foraging in schools, has probably been a major selective force promoting colonial nesting in centrarchid fishes.

This hypothesis is supported by the observation that those centrarchids which frequently nest solitarily or in loose association (for example, *Micropterus*, *Ambloplites* and *Pomoxis*) are larger, more piscivorous fishes than those that typically nest colonially, such as *Lepomis macrochirus*, *Lepomis megalotis* or *Lepomis punctatus*, even in the same habitat. These larger, more piscivorous forms are probably able to defend their nests better against foraging schools¹⁵. Such superior capability alone, however, is an inadequate defence for isolated nests when sunfish are present in high densities. In these conditions the nesting success of even the largest, most piscivorous forms, such as *Micropterus*, can be severely limited¹⁶. Finally, as foraging schools of bluegills are such predominant and effective predators on the eggs of conspecifics, wherever bluegills occur and

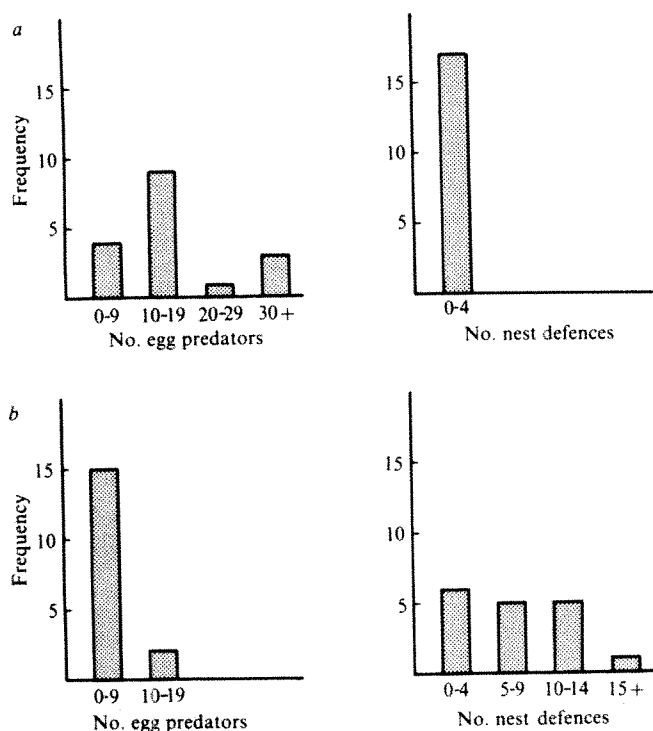


Fig. 2 Two observers simultaneously removed nest-guarding males from *a*, isolated nests (having no more than one neighbouring nest) and *b*, surrounded nests (having five or more neighbouring nests), then recorded the number of egg predators present in each nest at 1-min intervals for 10 min, and the total number of defences by neighbouring males. Surrounded nests had significantly fewer cumulative egg predators ($P < 0.001$, Mann-Whitney) and received significantly more defences by neighbouring males ($P < 0.001$, Mann-Whitney).

nest, an effective egg predator will also occur. Thus bluegills might be considered their own worst enemies, and may themselves have generated much of the selective pressure promoting their extreme form of colonial nesting.

I thank R. E. Buskirk, P. W. Colgan, S. T. Emlen, R. Hegner, S. Nowicki, P. W. Sherman, P. Wrege and J. Wimer for criticisms of the manuscript, the administration of Cazenovia College, J. Brady and D. Roy, Director of Cazenovia College Lake Research Station, for use of the facilities, and L. Olney and E. B. Brothers for equipment. A Donjacour, A. Horn, S. Letovsky, C. Mercaldi, A. Plassman, C. Van Buren, and J. Weiskotten were field assistants. Figure 1 was drawn by A. Plassman. The research was supported by a doctoral dissertation improvement grant from the NSF and by grants from the Ithaca and National chapters of Sigma Xi, the George D. Harris Foundation and the Theodore Roosevelt Memorial Fund of the American Museum of Natural History.

Received 15 September; accepted 31 December 1980.

- Kruuk, H. *Behav. Suppl.* **11**, 1-129 (1964).
- Tinbergen, N. *Ibis* **98**, 401-411 (1956).
- Horn, H. S. *Ecology* **49**, 682-694 (1968).
- Ward, P. & Zahavi, A. *Ibis* **115**, 517-534 (1973).
- Hoogland, J. L. & Sherman, P. W. *Ecol. Monogr.* **46**, 33-58 (1976).
- Darling, F. *Bird Flocks and the Breeding Cycle* (Cambridge University Press, London, 1938).
- Emlen, S. T. & Demong, N. J. *Science* **188**, 1029-1031 (1975).
- Williams, A. J. *Ornis Scandinavica* **6**, 117-124 (1975).
- Colgan, P. W., Nowell, W. A., Gross, M. R. & Grant, J. W. A. *Envir. Biol. Fish.* **4**, 29-34 (1979).
- Gerald, J. W. thesis, Univ. Texas (1970).
- Loiselle, P. V. *Biol. Behav.* **2**, 129-142 (1977).
- Dominey, W. J. *Nature* **284**, 546-548 (1980).
- Dominey, W. J. *Envir. Biol. Fishes* (in the press).
- Hamilton, W. D. J. *theor. Biol.* **31**, 295-311 (1971).
- Williams, G. C. *Adaptation and Natural Selection* (Princeton University Press, 1966).
- Smith, S. L. *Trans. Am. Fish. Soc.* **105**, 682-689 (1976).

Induction of subacute murine measles encephalitis by monoclonal antibody to virus haemagglutinin

Kottil W. Rammohan, Henry F. McFarland & Dale E. McFarlin

Neuroimmunology Branch, National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20205, USA

Although the events which predispose a host to measles virus persistence remain largely unknown, measles antibody has been shown to contribute to the production of a persistent infection by this virus both *in vivo* and *in vitro*. Thus, the addition of measles antibody to cells infected with measles virus promotes virus persistence^{1,2}. Latent measles infection occurs in newborn hamsters with maternally acquired antibody after inoculation with measles virus in the immediate neonatal period³. A subacute encephalitis with measles virus persistence has been induced in weanling BALB/c mice that received antibody after virus inoculation⁴ and in measles-immune primates infected with a virus derived from a patient with subacute sclerosing panencephalitis⁵. Measles virus consists of six polypeptides, and treatment of measles-infected cells with antibody has been shown⁶ to alter the pattern of their synthesis. As the antigenic specificity of the antibody responsible for these observations is not known, we decided to investigate the effects of monoclonal antibodies directed against the individual measles polypeptides. We report here that a monoclonal antibody directed against the virus haemagglutinin, unlike an antibody to the virus nucleocapsid protein, is able to induce a subacute encephalitis *in vivo*.

After the intracerebral (IC) inoculation of weanling BALB/c mice with $10^{2.7}$ ICD₅₀ units of hamster neurotropic (HNT) measles, about 70-80% of the mice die from an acute encephalopathy. This encephalopathy produces mortality within the first 3 weeks and is characterized by a complete absence of brain inflammation, although neuronal viral antigen is readily demonstrable⁷. If hyperimmune measles antiserum (HI serum) is administered 3 days after virus inoculation, all the mice survive the acute period, but ~30% develop a delayed disease and die. Its onset is variable and has been observed up to 8 weeks post-inoculation. It is characterized by weight loss, seizures, and multiple neurological abnormalities that are fatal in a few weeks. Unlike the acute encephalopathy where brain inflammation is absent, histopathology of this subacute disease is characterized by mild inflammatory response of the meninges and parenchyma. Thus, it is referred to as an encephalitis. Immunofluorescence of the frozen sections from moribund animals shows abundant viral antigen in neurones of the hippocampus, frontal cortex, brainstem and cerebellum⁴. Mice which do not develop the subacute disease (~70%) have been observed for up to 6 months and remain well.

The HI serum used to induce the subacute encephalitis was prepared by hyperimmunization of BALB/c mice with Edmonston measles virus⁴. The serum has reactivity to all polypeptides of the measles virion. However, the polypeptide specificities of the antibody molecules in this panreactive serum that are responsible for this *in vivo* effect are unknown. To investigate this question, ascitic fluid containing monoclonal antibodies with single specificities against the virus haemagglutinin (C1 AF)⁸ and virus nucleocapsid protein⁹ was tested in this system for the ability to modify disease. Administration of the C1 AF abolished the acute encephalopathy and led to the development of a subacute disease which lasted up to 70 days. Two representative experiments, out of eight, are shown in Table 1. Large doses of C1 AF inhibited disease completely and with decreasing amounts of antibody the incidence of the

Table 1 Mortalities of BALB/c mice after intracerebral inoculation of HNT measles virus—effects of passive transfer of monoclonal antibody to viral haemagglutinin

Expt no.	Group	Antibody	Dilution	Mortality		Mean survival time (days)	Daily mortalities
				Acute	Delayed		
1	A	None	—	70 (7/10)	0	11 (8–19)	d8 ³ , d9 ² , d16 ¹ , d19 ¹
	B	HI serum	1:10	0	50 (8/16)	39 (24–56)	d24 ¹ , d27 ¹ , d38 ³ , d41 ¹ , d48 ¹ , d56 ¹
	C	P3 X63 AF	1:10	70 (7/10)	0	14 (12–18)	d12 ² , d13 ³ , d17 ¹ , d18 ¹
	D	79 X1 C1 AF	1:25	0	30 (3/10)	56 (26–71)	d26 ¹ , d70 ¹ , d71 ¹
			1:100	0	40 (4/10)	43 (35–56)	d35 ¹ , d38 ¹ , d42 ¹ , d56 ¹
2	A	None	—	72 (8/11)	0	12 (9–19)	d9 ¹ , d10 ² , d11 ² , d12 ¹ , d14 ¹ , d19 ¹
	B	79 X1 C1 AF	1:10	0	0	—	—
			1:20	0	20 (2/10)	46 (36–56)	d36 ¹ , d56 ¹
			1:100	0	36 (9/25)	37 (28–45)	d28 ² , d32 ² , d41 ³ , d43 ¹ , d45 ¹
			1:500	10 (1/10)	40 (4/10)	25 (20–34)	d20 ¹ , d22 ² , d25 ¹ , d34 ¹

Four-week-old female BALB/c mice were inoculated with $10^{2.7}$ ICLD₅₀ units of HNT measles virus in 0.03 ml in the right cerebral hemisphere under light ether anaesthesia, using a tuberculin syringe and 30 G needle. All animals received 0.2 ml of the appropriate dilution of antiserum through the tail vein on day 3 post-inoculation. Control animals received nothing. Transfer of normal mouse serum in previous experiments were without effect on the onset or course of clinical disease. Characteristics of antisera used were: HI serum, serum pooled from 50 female BALB/c mice after hyperimmunization with purified Edmonston measles virus. Initial immunization was with complete Freund's adjuvant followed by multiple boosts (13 boosts over 18 months) in saline or incomplete Freund's adjuvant. The serum titres were: 1:1,024 by haemagglutinin inhibition test, 1:2,048 by complement fixation test, $>10^{-4}$ by virus neutralization test and 10^{-6} dilution by radioimmunoassay. The serum reacts with all polypeptides of the Edmonston measles virus in a radioimmunoprecipitation assay³. PS X63 AF, ascitic fluid prepared from pristane-primed BALB/c mice inoculated with the parent myeloma cells used to produce the 79 XI-C hybridomas⁶. 79 XI C1 AF, ascitic fluid prepared from pristane-primed BALB/c mice inoculated with 79 XI-C hybridoma. The ascitic fluid had reactivity only with the haemagglutinin of Edmonston measles virus. The ascitic fluid titres were: 1:4,000 dilution in virus neutralization, 1:4,000 in haemagglutinin inhibition and 10^{-7} dilution in a solid-phase radioimmunoassay. The antihaemagglutinin antibody is identified as IgG1 subclass⁶. All mortalities were recorded from the day of virus inoculation. Mortalities within the first 21 days were arbitrarily considered 'acute' and subsequent mortalities were included in the 'delayed' group. Mean survival time was calculated as the mean duration from the day of inoculation to the day of death (range indicated in parentheses). Daily mortalities; superscript number indicates the total mortality for the day (d) indicated.

subacute clinical disease increased. At the greatest dilution (1:500) used, protection from the acute disease was not absolute (expt 2, Table 1). In animals receiving the lower doses of antibody the onset of the subacute illness was earlier and the total mortality greater. These observations suggested an inverse correlation between the dose of antibody transferred and total mortality, and, furthermore, that the onset of clinical illness may be related to the fall in serum antibody. This was supported by studies correlating the antibody levels with the onset of clinical illness. Following transfer, the antibody levels declined, and were barely detectable when clinical illness began. Transfer of large amounts of monoclonal antibody to the virus nucleocapsid protein did not affect the acute mortalities in three experiments. Similarly, no effect was observed with normal mouse serum or ascitic fluid obtained from the parent myeloma (P3 X63)⁸.

The virological and pathological features of the delayed encephalitis were studied. One hundred weanling BALB/c mice were inoculated with $10^{2.7}$ ICLD₅₀ units of HNT measles virus and given 0.2 ml of a 1:100 dilution of the C1 AF through the tail vein on day 3. Every 5 days viral expression in the brains of five infected animals was studied by direct immunofluorescence. A fluorescein-conjugated IgG purified from the serum of a patient with subacute sclerosing panencephalitis capable of reacting with all measles polypeptides was used¹⁰. On day 8 after virus inoculation, the brains of all five animals had viral antigen, which was detectable in occasional neurones. On days 13 and 18 viral antigen could no longer be detected in any of the 10 brains studied. At day 23 viral fluorescence was readily observed bilaterally in all five animals. The onset of clinical illness followed shortly afterwards in the remaining mice. The animals with clinical encephalitis had viral antigen detectable in the neurones of the hippocampus, frontal cortex, brain stem and cerebellum. Brains from moribund animals were also studied by light microscopy. These showed meningeal and parenchymal perivascular cuffs of mononuclear cells. However, the degree of inflammation was considerably less than in chronic encephalitis described in the SJL mouse¹⁰ or disease induced by the panreactive serum in BALB/c mice⁴.

Clearly, unlike the antibody to nucleocapsid protein, antibody with reactivity only to the virus haemagglutinin can modify the clinical course of acute murine measles encephalitis. Although

the precise mechanisms whereby C1 AF delays the acute encephalopathy are unknown, the effect may be through antigenic modulation *in vivo*. The haemagglutinin is expressed on infected cell surfaces and is available for modulation by antibody. The observed disappearance of all viral antigens from the central nervous system at days 13 and 18 suggests that the reaction between the haemagglutinin peptide and the anti-haemagglutinin antibody alters the synthesis of all viral polypeptides *in vivo*. Similar effects have been observed in experiments *in vitro*⁶.

The HI serum contains antibody against the fusion protein which is also expressed on infected cell surfaces. Although antibodies to fusion protein may also be important in antigenic modulation of infected cells, *in vitro* studies suggest that this antigen is available for interaction with its antibody only in the Fab monomeric state at least on the surface of some infected cells¹¹. However, monospecific antibody to the fusion protein of another paramyxovirus, SV5, inhibited the spread of infection by limiting cell fusion and virus entry after virus adsorption¹². The effect of antibody to the fusion protein on the modulating ability of the HI serum remains to be determined. The results from this study establish that antibody to one surface component, the haemagglutinin, is responsible for some of the *in vivo* effects of the HI serum. The antibody reactivity not only leads to delayed clinical disease but alters the synthesis of viral polypeptides in infected cells.

We thank Drs Barry Bloom and M. Birrer for providing the hybridoma cells with antinucleocapsid reactivity, and Mrs Shirley Sutphin and Gail Raoofi for secretarial assistance.

Received 10 November 1980; accepted 16 February 1981.

1. Rustigian, R. J. *Bact.* **92**, 1792–1804 (1966).
2. Joseph, B. J. & Oldstone, M. B. A. *J. exp. Med.* **142**, 864–876 (1975).
3. Wear, D. J. & Rapp, F. J. *Immun.* **107**, 1593–1598 (1971).
4. Rammohan, K. W., McFarland, H. F. & McFarlin, D. E. *New Engl. J. Med.* (submitted).
5. Albrecht, P., Burnstein, T., Klutch, M. J., Hicks, J. T. & Ennis, F. A. *Science* **195**, 64–66 (1977).
6. Fujinami, R. S. & Oldstone, M. B. A. *Nature* **279**, 529–530 (1979).
7. Griffin, D. E., Mullinix, J., Narayan, O. & Johnson, R. T. *Infect. Immun.* **9**, 690–695 (1974).
8. McFarlin, D. E., Bellini, W. J., Mingioli, E. S., Behar, T. N. & Trudgett, A. J. *gen. Virol.* **48**, 425–429 (1980).
9. Birrer, M. J., Bloom, B. R. & Udem, S. *Virology* (in the press).
10. Rammohan, K. W., McFarlin, D. E. & McFarland, H. F. *J. infect. Dis.* **142**, 608–613 (1980).
11. Sissons, J. G., Cooper, N. R. & Oldstone, M. B. A. *J. Immun.* **123**, 2144–2149 (1979).
12. Merz, D. C., Scheid, A. & Chopin, P. W. *J. exp. Med.* **151**, 275–288 (1980).

Acute rabies death mediated by antibody

Bellur S. Prabhakar* & Neal Nathanson

Department of Microbiology, University of Pennsylvania Medical School, Philadelphia, Pennsylvania 19104, USA

Among inadequately immunized laboratory animals dying of rabies, a small but consistent proportion succumb after an incubation period shorter than that of any of the unvaccinated controls^{1-3,7} and this phenomenon has been termed 'early death'. It has also been shown that, after immunosuppression either with cyclophosphamide or by thymectomy and irradiation⁴, rabies-infected animals survive longer with decreased incidence of paralysis. It appears that 'early death' also occurs in humans who have been treated (unsuccessfully) with vaccine, with or without accompanying serum therapy⁵⁻⁷, after exposure to rabies. Collectively, these studies suggest that there is an immunopathological component in rabies virus infection. We have therefore investigated a model of lethal rabies infection in immunosuppressed mice and have concluded that B lymphocytes or antibody specific for rabies play a part in the causation of early death. Accordingly the immune response to rabies has a dual role⁸, sometimes favouring survival but sometimes enhancing the disease.

In this study, we used the high-egg-passage Flury (HEP) strain of rabies virus, which causes a limited inapparent infection of the central nervous system (CNS) in adult mice⁹ that can be converted into a fatal infection by immunosuppression¹⁰. This fatal infection can in time be aborted by adoptive transfer of syngeneic immune spleen cells¹¹. Adult BALB/c mice were lethally irradiated by total body irradiation (TBI) from a ¹³⁷Ce source (800–850 rad). TBI and normal mice were infected intracerebrally (IC) on the day of irradiation with 5,000 suckling mouse ICLD₅₀ of HEP virus. As shown in Table 1, untreated TBI mice died due to radiation; but could be protected by normal bone marrow cells devoid of mature T lymphocytes. Normal mice infected with HEP virus survived, but HEP-infected TBI mice died despite immunological reconstitution with bone marrow cells, clearly separating rabies death from that due to radiation.

HEP-infected TBI mice reconstituted with normal spleen cells or cells from donors immunized 30 days earlier (memory cells), died after a mean survival time (MST) of 9 days, but spleen cells from donors immunized 7 days earlier (primary immune cells) conferred complete protection with animals showing no apparent symptoms of the disease.

However, adoptive transfer of spleen cells obtained from donors 3 days after secondary immunization (secondary immune cells) resulted in early death of 28% of the recipients, while the rest survived. The animals that died within 2–3 days after cell transfer did not show any signs of paralysis, typical of rabies, but rather acute symptoms such as convulsions and tonic spasms. Onset of these symptoms was noted only a few hours before death. Histopathological examination of parasagittal sections of the brain from nine of these animals that were either moribund or just dead showed no inflammation, neuronal destruction or other alteration. By contrast, in brains obtained immediately after death from a group of eight mice dying of rabies 13–15 days after infection, foci of necrotic hippocampal cells with minimal inflammation were consistently seen.

To characterize further the immunological requirements for early death, secondary immune spleen cells from syngeneic donors were separated into B- and T-cell enriched populations by passage through a nylon wool column¹² or by treatment with anti- θ antibody and complement. TBI mice infected with HEP

Table 1 Effect of TBI and adoptive cell transfer on HEP virus infections

TBI	HEP	Donor cells	Mortality		
			Early death*	Total	MST
+	—	—	0/10	10/10	9
+	—	Bone marrow	0/15	0/15	—
—	+	—	0/10	0/10	—
+	+	—	0/10	10/10	8
+	+	Bone marrow	0/20	20/20	10
+	+	Normal spleen	0/14	14/14	9
+	+	Immune spleen			
		1° + 7†	0/20	0/20	—
		1° + 30	0/12	12/12	9
		2° + 3	9/32	9/32	3

10–12-week old BALB/c mice were irradiated (800–850 rad, TBI) and on the same day were inoculated with 5,000 suckling mouse ICLD₅₀ of HEP virus. Four days later, these mice were injected intravenously (i.v.) with 10⁷ bone marrow cells or 10⁸ spleen cells from syngeneic donors. MST, mean survival time (days) for mice that died, counting from the 4th day of infection when cells were transferred.

* Indicates mortality within 72 h of cell transfer.

† Indicates primary or secondary immune cells and the time of collection, in days, after last donor immunization.

Table 2 Induction of antibody-mediated death in TBI and HEP virus-infected mice

Immunotherapy	Early death* (%)	Total mortality (%)
B cells	6/27 (22)	100
T cells	0/24 (0)	100
Anti-rabies antibody	6/24 (25)	100
Normal serum	0/12 (0)	100

10–12-week old BALB/c mice were irradiated (800–850 rad) and, on the same day, inoculated with 5,000 suckling mouse ICLD₅₀ of HEP virus. T and B cells were obtained from donors immunized with rabies virus 33 and 3 d before collection of spleen cells; 60 × 10⁶ fractionated viable cells were given i.v. on day 4. Mouse hyperimmune anti-rabies antibody and normal serum were given intraperitoneally in volumes of 0.5 ml each on day 4.

* Death within 72 h of cell or serum transfers.

were given enriched B cells, T cells, mouse anti-rabies antibody or normal mouse serum. The data summarized in Table 2 show that either immune B lymphocytes or rabies-specific antibody, similar to secondary immune cells, can induce early death, whereas T lymphocytes failed to do so. The failure of B or T lymphocytes or antibody to protect is probably due to incomplete immunological reconstitution of these animals^{11,13–15}.

These results clearly demonstrate that early death from rabies virus is mediated by antibody rather than by immune T cells. However, the mechanisms by which the antibody precipitates the disease are not obvious. Although it is well established that anti-rabies antibody plus complement can lyse rabies-infected cells in culture¹⁶, involvement of such a mechanism in early CNS lesions needs further study. This model may lead to further understanding of the contribution of host immunological responses to rabies pathogenesis, and may also be relevant to the pathogenesis of obscure virus-initiated human diseases such as the acutely fatal encephalopathy of Reye's syndrome.

We thank Dr T. J. Wiktor for allowing this work to be done in his laboratory at the Wistar Institute, and Drs G. A. Cole and J. R. Klein for their critical comments. This work was supported in part by USPHS grants NS 16010, AI 09706 and CA 09410.

Received 22 December 1980; accepted 9 February 1981.

1. Sikes, R. K., Cleary, W. F., Koprowski, H., Wiktor, T. J. & Kaplan, M. M. *Bull. Wild Hlth Org.* **45**, 1–11 (1971).
2. Baer, G. M. & Cleary, W. F. *J. infect. Dis.* **125**, 520–527 (1972).
3. Veeraraghavan, J. *Ann. Rep. of the Director* (The Pasteur Institute of Southern India, Conoor, 1971).

* Present address: Laboratory of Oral Medicine, NIDR, National Institutes of Health, Building 30, Room 121, Bethesda, Maryland 20205, USA.

4. Tignor, G. H., Shope, R. E., Gershon, R. K. & Laksman, B. H. *J. Immun.* **112**, 260–265 (1974).
5. Anonymous 1969–1975 *Vigilância Epidemiológica* **8**, No. 9 (Oficina Sanitaria Pamericana, 1976).
6. Kyavchenko, A. T., Karerkuymchan, M. K. & Wsenko, A. L. *Zh. Mikrobiol. Épidem. Immunobiol.* **12**, 44–51 (1974).
7. Held, J. R., Tierkel, E. S. & Steele, J. H. *Publ. Hlth Rep.* **82**, 1009–1018 (1967).
8. Nathanson, N., Johnson, E. D., Camenga, D. L. & Cole, G. A. in *4th int. Convoc. Immun.* (eds Neter, E. & Milgrom, F.) 76–94 (S. Karger, Basel, 1974).
9. Koprowski, H., Black, J. & Nelson, D. J. *J. Immun.* **72**, 94–106 (1954).
10. Fischman, H. R. & Strandberg, J. D. *J. Am. vet. med. Assoc.* **163**, 1050–1055 (1973).
11. Prabhakar, B. S., Nathanson, N. & Wiktor, T. J. *Fedn Proc.* **39**, 355 (1980).
12. Julius, J. H., Simpson, E. & Harzenberg, L. A. *Eur. J. Immun.* **3**, 645–649 (1973).
13. Miller, A., Morse, H. C. III., Winkelstein, J. & Nathanson, N. *J. Immun.* **121**, 321–326 (1978).
14. Kaplan, M. M., Wiktor, T. J. & Koprowski, H. *J. Immun.* **114**, 1761–1765 (1975).
15. Turner, G. S. *J. gen. Virol.* **33**, 535–538 (1976).
16. Wiktor, T. J. & Clark, H. F. *Infect. Immunity* **6**, 988–995 (1972).
17. Blancou, J., Andral, B. & Andral, L. *J. gen. Virol.* **50**, 433–435 (1980).

Monoclonal antibody identifies a new Ia-like (p29,34) polymorphic system linked to the HLA-D/DR region

Lee M. Nadler*, Philip Stashenko†, Russell Hardy*, Kevin J. Tomaselli*, Edmond J. Yunis‡, Stuart F. Schlossman* & John M. Pesando*

The Divisions of Tumor Immunology* and Immunogenetics‡, Sidney Farber Cancer Institute–Harvard Medical School, Boston, Massachusetts 02115, USA

Department of Immunology†, Forsyth Dental Center–Harvard School of Dental Medicine, Boston, Massachusetts 02115, USA

The Ia antigens of the mouse are the basis for the genetic control of the immune response. The HLA-D/DR locus is considered to be the human counterpart of the Ia subregion of the murine major histocompatibility complex^{1–4}. The HLA-D/DR antigens are polymorphic, and eight well defined alleles have been identified using alloantisera⁵. More recently, 'supertypic' antigens (MB and MT) have been defined which identify clusters of HLA-D/DR specificities^{6–9}. Little is known about the molecular basis for the cellular and serological polymorphism of the HLA-D/DR antigens, as alloantisera are usually of very low titre and heteroantisera frequently lack monospecificity. We present here the preparation and characterization of a monoclonal antibody which defines a new polymorphic system of the HLA-D/DR region. This and similar antisera should now begin to provide the reagents with which to correlate molecular structure with the functional repertoire of the human Ia-like antigens.

Six-week old BALB/c mice were immunized intraperitoneally with 5×10^6 ascites tumour cells from a patient with Burkitt's lymphoma. Four weeks later, a mouse was boosted intravenously with 5×10^6 tumour cells and somatic cell hybridization was carried out 4 days later by the modified¹⁰ method of Kohler and Milstein¹¹. Splenocytes (1×10^8) were fused in 30% polyethylene glycol with P3/NS1/1-Ag4-1 myeloma cells (2×10^7) and aliquoted into six microtitre plates (Falcon). Aminopterin was added 24 h after fusion, and macroscopic hybridoma colonies became evident between 10 and 28 days in culture.

Preliminary screening identified several hybridoma antibodies with reactivity limited to: (1) the immunizing Burkitt's lymphoma tumour cell, (2) B-cell lines, but not T-cell lines, and (3) 20–25% of Ficoll–Hypaque fraction of peripheral blood mononuclear cells from normal individuals. All cell-surface characterization was done by standard indirect immunofluorescence assay with subsequent analysis of each population by the fluorescence-activated cell sorter (FACS-I, Becton-Dickinson)¹².

Antisera were then evaluated to determine whether their cellular reactivity was similar to that previously described for anti-Ia-like antisera. The peripheral blood of several individuals

was separated, as previously described¹³, into T, B, monocyte and null cell subpopulations. In addition, purified T cells were stimulated with phytohaemagglutinin (PHA) or concanavalin A (Con A). One antibody, designated I-LR1, reacted with >90% of B cells and monocytes, ~25% of null cells and 25% of Con A- or PHA-activated T cells, but was unreactive with freshly isolated or unstimulated cultured T cells. An identical pattern of reactivity with these fractionated cells was found with the previously reported¹⁴ I-1 monoclonal antibody which defines a non-polymorphic Ia-like antigen. In addition, antibodies I-1 and I-LR1 both identify bimolecular glycoprotein complexes composed of subunits of molecular weights 29,000 and 34,000 (see below and Fig. 1) which have been previously shown using heteroantisera to define human Ia-like antigens^{15,16}. Hybridoma I-LR1 was then cloned by limiting dilution and injected into pristane (Aldrich)-primed BALB/c mice to produce immune ascites¹⁷. Monoclonal antibodies I-LR1 and I-1 were then tested on the B cell-enriched mononuclear cells from 50 healthy

Table 1 Reactivity of I-LR1 with homozygous typing cells

Dw	DR	No. HTC's tested*	No. positive with monoclonal antibody	
			I-1	I-LR1
1	1	5	5	4
2	2	9	9	3
3	3	8	8	2
4	4	6	5	0
5	5	2	2	2
6	w6	3	3	2
7	7	5	5	2
10	4	3	3	0
11	7	6	6	3
Total		47	46	18

Cryopreserved HTC's were provided by the National Institute of Allergy and Infectious Disease HTC Typing Panel and by the Histocompatibility Laboratory of the Johns Hopkins Medical School. The NIAID HTC panel represented homozygous cells from several typing laboratories, including: (1) the Sidney Farber Cancer Institute, (2) Memorial Sloan-Kettering Cancer Center, (3) Stanford University, (4) Columbia University, and (5) Leiden. These cells were typed for HLA-A, -B, -C, -DR and -D, and are presented by D typing. They include: Dw1 (BVR, OOS, RM-8W104, CAH and J. Beil); Dw2 (FVV, VN2, KJL, TAL, TI, PA, NOL, LA and VML-8W113); Dw3 (AVL, NR3-8W122, NC-3, TPA-8W123, SFN RW, HRA, 2853-8W119 and KF-8); Dw4 (SFN BN, RJM, MWM, GW-8W132, ER-81128 and 5-0107); Dw5 (NAB-8W142 and 6W7006-8W137); Dw6 (HHK, JGr and Daudi); Dw7 (P. Burk, SFN MA, JB12 and MMF); Dw10 (EM10-8W210, FS-8W212 and 2046-8W213); and Dw11 (DMT, MSA, CA, JVM, J. King and D. Beil). (8W represents the Eighth Histocompatibility Workshop designation of the HTC's.) All analyses were carried out on cryopreserved cells with B cell-enriched fractions isolated by Sephadex G-200 anti-F(ab')₂ chromatography. Cells ($1-2 \times 10^6$) were treated with 0.1 ml of a 1:5,000 dilution of an anti-Ia monoclonal antibody or a 1:5,000 dilution of an unreactive control antibody of a similar immunoglobulin isotype, incubated at 4°C for 30 min, and washed twice. The cells were then reacted with 0.1 ml of a 1:40 dilution of a fluoresceinated goat anti-mouse IgG at 4°C for 30 min, washed twice, and analysed on the FACS-I or Cytofluorograf FC200/4800A. Intensity of fluorescence for I-1 or I-LR1 was determined for 40,000 cells in each population by standard indirect immunofluorescence and compared with the fluorescence of an unreactive isotype identical monoclonal antibody. A displacement of the histogram of the test monoclonal antibody compared with the histogram of an unreactive isotype monoclonal antibody was scored positive. In addition, for each test sample, a quantitative assessment of the number of positive cells was made (number of cells reactive with test monoclonal antibody minus number of cells reactive with the unreactive isotype-identical monoclonal antibody divided by 40,000 total cells tested).

* In addition to B cell-enriched fractions isolated from HTC mononuclear cells, 14 HTC's were transformed by Epstein-Barr virus into lymphoblastoid B-cell lines and the identical patterns of reactivity were found for I-1 and I-LR1 (two lines of each of the following D/DR types: 1,2,3,4,5,7,11).

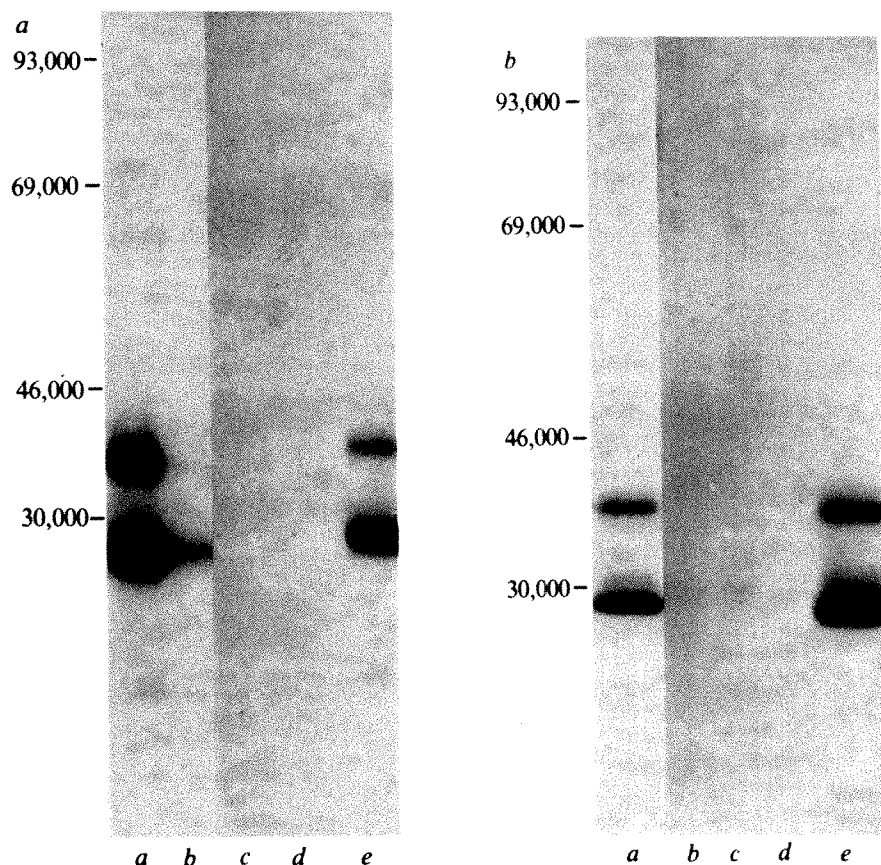


Fig. 1 Serial immune precipitation using monoclonal antibodies I-1 and I-LR1. **A:** *a*, I-1; *b*, I-1; *c*, I-1; *d*, protein A-Sepharose alone; *e*, I-LR1. **B:** *a*, I-LR1; *b*, I-LR1; *c*, I-LR1; *d*, protein A-Sepharose alone; *e*, I-1. The B-lymphoblastoid cell line, Laz 388, was surface-labelled with ^{125}I by the lactoperoxidase method²³. NP40-solubilized membrane extracts were applied to a lentil lectin column and bound glycoproteins eluted with 3% α -methyl-D-mannoside. In conditions of antibody excess, an aliquot of lentil lectin-retained labelled membrane material was incubated three times with I-1 bound to protein A-Sepharose for 14 h at 4 °C. After each extraction, pellets were washed three times in a detergent containing buffer and counted along with a control sample to follow the efficiency of the extraction process. The supernatant was then incubated for 14 h with protein A-Sepharose without added antibody and processed in a similar way to control for any free antibody-antigen complexes which might be present. The supernatant was then incubated overnight with I-LR1 bound to protein A-Sepharose. Similar serial extractions were performed with the I-LR1 antibody. Antigens were resolved by electrophoresis on a gradient polyacrylamide slab gel and visualized by fluorographic processing and autoradiography as described previously²³. Molecular weights are shown.

unrelated HLA heterozygous individuals. Antibody I-LR1 was reactive with the B cells of only 28 individuals, whereas antibody I-1 reacted with all 50, suggesting that the I-LR1 antigen was polymorphic.

To investigate the specificity of I-LR1 further, a panel of homozygous typing cells (HTCs) was examined. As shown in Table 1, I-1 reacted with the B cell-enriched fraction of 46 of 47 HTCs tested, and I-LR1 reacted with only 18 of 47. Moreover, the pattern of reactivity of I-LR1 did not identify a monospecific HLA-D/DR determinant, nor did it identify a previously defined cluster of HLA-D/DR determinants (supertypic). In addition, five separate pairs of HLA-D/DR-identical individuals (pairs including HLA-D/DR 2,3; 4,7; 4,6; and 1,4) displayed reciprocal patterns of reactivity with I-LR1 (for example, D/DR 2,3 reactive, I-LR1 reactive and D/DR 2,3 reactive, I-LR1 nonreactive). This pattern of reactivity on HTCs suggests that I-LR1 defines a new Ia-like specificity.

Additional support for the notion that the I-LR1 antigen was encoded by the major histocompatibility complex was demonstrated by family inheritance studies. As shown in Table 2

(family Kan), I-LR1 reactivity segregates with the maternal 'c' haplotype. This observation strongly suggests that the I-LR1 locus segregates with alleles linked to the HLA complex. Linkage to the HLA-D/DR locus was then demonstrated in families with HLA-D/DR recombinant individuals. As shown in Table 2 (family Sch), the I-LR1 reactivity segregates with the maternal 'c' haplotype. Children 3 and 4 are HLA identical, except at the HLA-D/DR region (see Table 2 legend). B cells from child 3 were reactive with I-LR1, whereas the B cells from child 4 were nonreactive, suggesting that the I-LR1 locus is closely linked to the HLA-D/DR locus. Three additional families (data not shown) provided informative linkage data for I-LR1. For the total of five families (26 of 26 individuals), a determination of the linkage to HLA-D/DR was determined by the lod score method¹⁸, and a very tight linkage to the HLA-D/DR was found ($\theta_{\text{max}} = 0.0$, lod score 6.321). These family studies demonstrate that I-LR1 is linked to the HLA-D/DR locus.

The present study demonstrates that the I-LR1 antibody identifies a polymorphic antigen which shares a pattern of

Table 2 Inheritance of I-LR1 within families

Family identity	Antibody reactivity	Paternal haplotypes $\left(\frac{a}{b}\right)$	Maternal haplotypes $\left(\frac{c}{d}\right)$	Reactivity with siblings							
				1	2	3	4	5	6	7	8
Kan	I-LR1	A1, Cw*, B8, Dw3, DR3	A3, Cw-, B18, Dw1, DR1	a	a	b	b	a	b	a	b
		A3, Cw-, B7, Dw1, DR1	A28, Cw4, Bw35, Dw2, DR2	c	d	d	d	c	d	d	d
		0	+	+	0	0	0	+	0	0	0
Sch	I-LR1	A1, Cw-, B8, Dw3, DR3	A2, Cw5, Bw44(B12), Dw5, DR5	a	b	a	a	-	-	d	v
		A23, Cw4, Bw44(B12), Dw7, DR7	A3, Cw6, B57(B17), Dw7, DR7	d	c	c	c	-	-	c	c
		0	+	0	+	+	0	+	+	0	0
Recombinant—Sch family		Child 3	Child 4—recombinant HLA-B/D								
		$\frac{a}{c}$	$\frac{a}{c/d}$								
		A1, Cw-, B8, Dw3, DR3	A1, Cw-, B8, Dw3, DR3								
		A2, Cw5, Bw44(B12), Dw5, DR5	A2, Cw5, Bw44(B12), Dw7, DR7								

* A dash indicates that no typing was assignable with present international typing reagents.

cellular reactivity, molecular weight and HLA linkage with those previously described for human Ia-like antigens. However, the pattern of reactivity on HTC and HLA-D/DR-identical heterozygotes suggests that this antibody defines a new Ia-like antigenic specificity distinct from the known DR, MB and MT specificities. Lampson and Levy¹⁹ have recently demonstrated that two monoclonal antibodies defining non-polymorphic Ia-like antigens identify distinct populations of Ia-like (p29,34) molecules. We now have evidence that the p29,34 molecule identified by I-LR1 is distinct from the p29,34 molecule defined by a monoclonal antibody, I-1 (ref. 14), directed against a non-polymorphic Ia-like antigen (Fig. 1). Extraction of radiolabelled membrane material from the B-lymphoblastoid cell line Laz 388 with I-1 monoclonal antibody removes all the p29,34 glycoproteins recognized by that reagent (Fig. 1A, a-e). When this extracted supernatant was tested with I-LR1, clearly defined gp29,34 complexes were still identified. Identical serial extraction using I-LR1 followed by testing with I-1, produced similar results (Fig. 1B, a-e). As complete extraction of the Ia-like complex recognized by I-1 leaves the Ia-like complex recognized by I-LR1 and vice versa, these antibodies seem to identify different Ia-like molecules which are immunologically as well as genetically distinct.

Using N-terminal amino acid sequencing, several laboratories^{20,21} have demonstrated that the DR and Ia-like antisera so far tested define antigens which share a high degree of sequence homology with murine I-E/C antigens. Recently, Goyert *et al.*²², using peptide mapping and limited N-terminal amino acid sequencing, have found that the heavy subunits of the DRw2 and DRw7 antigens are nearly identical but differ significantly from the respective subunit of the DRw5 antigen. They have, therefore, speculated that DRw5 may represent the human counterpart of the mouse I-A subregion antigen. It is intriguing to speculate that the I-LR1 antigen may identify one allele of a new polymorphic system which may be similarly encoded by the human counterpart of the murine I-A locus. The precise relationship of the I-LR1 antigen to the murine Ia antigens will require further biochemical investigations.

We thank Dr Devendra P. Dubey for helpful discussions, Mr John Daley for technical assistance, and Ms Luci M. Grappi for secretarial assistance. This work was supported by NIH grants AI 12069, CA 19589, CA 06516, CA 20531, DE 04881, RR 05526, and NIH contract NO1 AI 82550. L.M.N. is a recipient of a Research Fellowship from the Medical Foundation, Inc., Massachusetts, J.M.P. of a Fellowship from the Cancer Research Institute.

Received 13 August 1980; accepted 6 February 1981.

- Barnstable, C. J., Jones, E. A. & Crumpton, M. J. *Br. med. Bull.* **34**, 241-246 (1978).
- Bodmer, J. G. *Br. med. Bull.* **34**, 233-239 (1978).
- Giphart, M. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 3533-3536 (1977).
- Allison, J. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3953-3956 (1978).
- Dupont, B., Braun, D. W., Yunis, E. J. & Carpenter, C. B. in *Histocompatibility Testing, 1980* (ed. Terasaki, P. I.) (UCLA Press, Los Angeles, in the press).
- Duquesnoy, R., Marrai, M. & Vieira, J. in *Histocompatibility Testing, 1980* (ed. Terasaki, P. I.) (UCLA Press, Los Angeles, in the press).
- Park, M. D., Terasaki, P. I., Nakata, S. & Aoki, D. in *Histocompatibility Testing, 1980* (ed. Terasaki, P. I.) (UCLA Press, Los Angeles, in the press).
- Brodsky, F. M., Parham, P., Barnstable, C. J., Crumpton, M. J. & Bodmer, W. F. *Immun. Rev.* **47**, 3-61 (1979).
- Trucco, M. M., Garotta, G., Stocker, J. W. & Ceppellini, R. *Immun. Rev.* **47**, 219-252 (1979).
- Kennett, R. H., Denis, K. A., Tung, A. S. & Klinman, N. R. *Curr. Topics Microbiol. Immun.* **81**, 77-91 (1978).
- Köhler, G. & Milstein, C. *Nature* **256**, 495-497 (1975).
- Nadler, L. M., Stashenko, P., Hardy, R. & Schlossman, S. F. *J. Immun.* **125**, 570-577 (1980).
- Chess, L., MacDermott, R. P. & Schlossman, S. F. *J. Immun.* **113**, 1113-1121 (1974).
- Reinherz, E. L. *et al. J. exp. Med.* **150**, 1472-1482 (1979).
- Crumpton, M. J. *et al. Proc. R. Soc. B202*, 159-175 (1978).
- Humphreys, R. E. *et al. J. exp. Med.* **144**, 98-112 (1976).
- Stashenko, P., Nadler, L. M., Hardy, R. & Schlossman, S. F. *J. Immun.* **125**, 1678-1685 (1980).
- Morton, N. E. *Am. J. hum. Genet.* **7**, 277-318 (1955).
- Lampson, L. A. & Levy, R. *J. Immun.* **125**, 293-299 (1980).
- Uhr, J. W., Capra, J. D., Vitetta, E. S. & Cook, R. G. *Science* **206**, 292-297 (1979).
- Silver, J., Walker, L. E., Reisfield, R. A., Pellegrino, M. A. & Ferrone, S. *Molec. Immun.* **16**, 37-42 (1979).
- Goyert, S. M., Hubert, J. J., Curry, R. A. & Silver, J. *Hum. Immun.* **1**, 161-175 (1980).
- Pesando, J. M. *et al. J. Immun.* **124**, 2794-2799 (1980).

A cyanine dye distinguishes between cycling and non-cycling fibroblasts

R. L. Cohen*, K. A. Muirhead*, J. E. Gill*†, A. S. Waggoner‡ & P. K. Horan*§

* University of Rochester Medical Center, Department of Pathology, PO Box 626, 601 Elmwood Avenue, Rochester, New York 14642, USA

† Department of Chemistry, Amherst College, Amherst, Massachusetts 01002, USA

Cellular proliferative activity has previously been determined by measuring the incorporation of radiolabelled nucleotides or by visual inspection of cellular morphology. Although two flow cytometric methods^{1,2} have recently been developed which can distinguish cycling from non-cycling cells, both have serious disadvantages. One method requires uptake of a substantial amount of BUdR¹, limiting its usefulness for *in vitro* systems. The other method utilizes RNA/DNA content differences² but its successful application has proved cell-type dependent. We have now used the findings that the cell membrane is more highly polarized in resting than in proliferating cells³ and that cyanine dyes carrying a delocalized positive charge enter live cells to an extent that depends on the cell membrane potential⁴, to develop a method of distinguishing between cycling and non-cycling cells. The greater the membrane polarization, the greater is the concentration of dye within the cell. At high concentrations, the dye molecules aggregate and their fluorescence is quenched⁵. Thus, for a given external dye concentration, cells of different membrane potential would accumulate different amounts of fluorescent (non-aggregated) dye. Using fibroblasts in culture conditions chosen to provide various models of cycling and non-cycling cells, we found that fluorescence intensity with the dye 3,3'-diheptyloxycarbocyanine (Di-O-C₇(3))⁶ was consistently greater in the former than the latter.

The human foreskin fibroblasts used in these experiments show virtually no mitotic activity once confluent. In the first

Table 1 Changes in fluorescence intensity of cycling and non-cycling fibroblasts with dye concentration

Dye concentration (M)*	Wound margin†	Confluent region†	Ratio‡	Background§
2.5×10^{-6}	246.0 ± 9.40	183.0 ± 12.7	1.34 ± 0.11	4.80 ± 0.32
1.0×10^{-6}	290.0 ± 12.1	112.0 ± 6.7	2.59 ± 0.35	5.50 ± 0.30
5.0×10^{-7}	258.0 ± 16.5	46.8 ± 3.1	5.51 ± 0.51	2.55 ± 0.03
2.5×10^{-7}	104.0 ± 7.50	22.9 ± 2.2	4.56 ± 0.54	2.62 ± 0.15
1.0×10^{-7}	70.7 ± 8.90	13.3 ± 1.9	5.32 ± 1.01	1.93 ± 0.16
5.0×10^{-8}	21.1 ± 1.34	8.34 ± 0.9	2.53 ± 0.42	1.68 ± 0.20

Slides were analysed with a Zeiss Axiomat photometer. Fluorescence was excited with epi-illumination from a 100-W mercury arc lamp; the exciter filter was a BP 450-497, the dichroic mirror an FT 510 and the barrier filter an LP 526. Areas of interest on the slide were initially identified with a $\times 10$ or $\times 25$ objective under bright-field phase. Measurements of cytoplasmic fluorescence from selected cells were made with a $\times 50$ 0.95 NA dry objective and an effective photometer aperture of 4- μ m diameter. Fluorescence collected by this aperture was detected by a Hamamatsu R-928 photomultiplier and the output pulses from the tube were detected and counted by a Princeton Applied Research Corporation Model 1105/1120 ratemeter/discriminator. Fluorescence intensity measurements were made immediately after opening the fluorescence excitation shutter. Although some fading or enhancement of fluorescence was observed over a 30-s period, the rate of change was sufficiently low to permit reproducible measurement of fluorescence intensity, as indicated by the standard deviations.

* Di-O-C₇(3); see Fig. 1 legend for a description of the staining method.

† Thousands of photon counts per cell; expressed as mean $\pm$ s.e. ($n = 15$) after subtraction of indicated background.

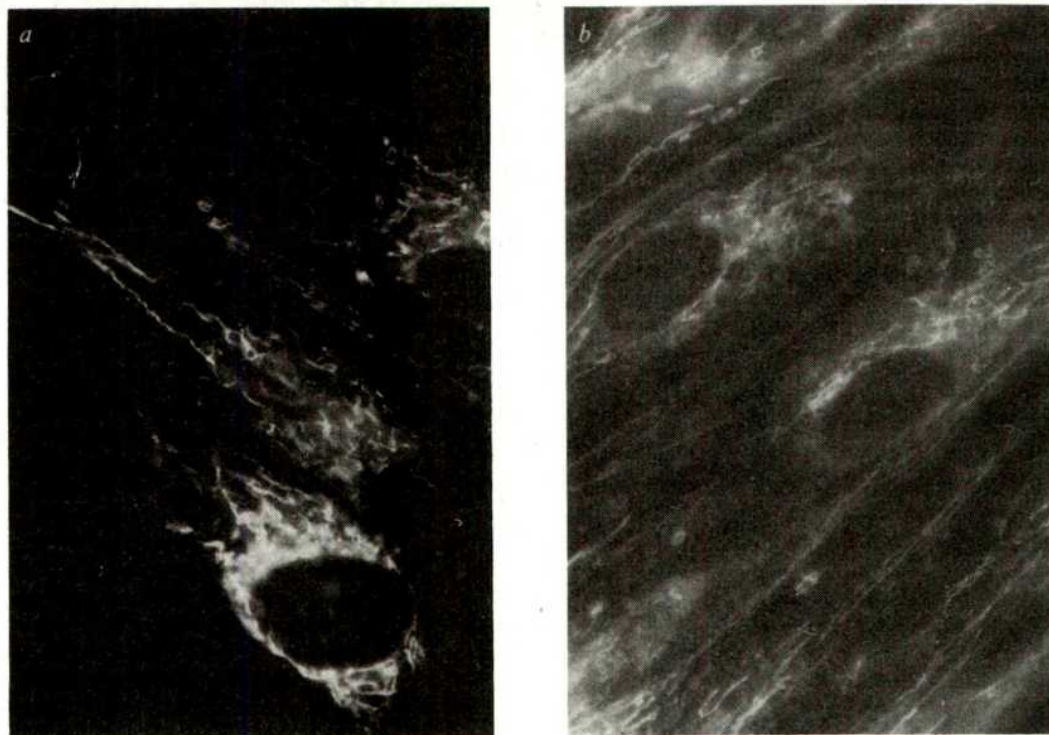
‡ Wound margin mean/confluent region mean.

§ Thousands of photon counts as measured in the centre of the wound away from all cells; expressed as mean $\pm$ s.e. ($n = 4$).

† Present address: FACS Division, Becton-Dickinson, 490-B Lakeside Drive, Sunnyvale, California 94086, USA.

§ To whom correspondence should be addressed.

Fig. 1 Fluorescence distribution in fibroblasts stained with the cyanine dye, Di-O-C₇(3). Human foreskin fibroblasts were wounded and stained with 1.0×10^{-7} M Di-O-C₇(3) 4 h after wounding. Fluorescence photomicrographs were taken using a $\times 100$ oil immersion objective and identical exposure times for cells at the wound margin (a) and cells in the confluent region of the same coverslip (b). Maximal definition of the fibrillar structures shown was observed at a stain concentration of 1.0×10^{-7} M. At 5×10^{-7} M more diffuse cytoplasmic staining was observed, although the intensity ratio for wounded and confluent cells was not significantly different at the two concentrations (Table 1). Di-O-C₇(3) (synthesized by A.S.W.) exhibits maximum absorbance at 488 nm (in ethanol) and maximum fluorescence at 504 nm. Stock dye solutions were prepared by adding ~ 1 mg of crystalline dye per ml of ethanol. Stock concentrations were assessed spectrophotometrically, based on an extinction coefficient of 1.49×10^4 M⁻¹ cm⁻¹. Stock solutions were stored for up to 2 months in foil-wrapped tubes at 4 °C. The human skin fibroblasts (donor HS0001, Naval Bioscience Laboratory, Oakland) were used for passages 1–5 and grown at 37 °C on coverslips placed in sealed Leighton tubes containing 10% CO₂. The growth medium consisted of DMEM supplemented with 10% fetal calf serum, penicillin (0.2 units ml⁻¹) and streptomycin (0.2 µg ml⁻¹). Growth medium was changed every 3–4 days. Wounding of monolayers was accomplished by scraping a strip of confluent cells from a coverslip using a modified metal microspatula and sterile procedures. Tubes containing scraped coverslips were returned to the incubator until staining. The cells were stained using the following procedure. Culture tubes were removed from the incubator and the growth medium was removed by aspiration. The slides within the tubes were then rinsed twice before staining. The rinse solution was Dulbecco's phosphate-buffered saline supplemented with 4.5 g l⁻¹ glucose. The staining solution consisted of the above salt solution plus an appropriate concentration of dye. The dye was allowed to equilibrate with the cells for 15 min at room temperature in the dark. Coverslips were then removed from the tubes with forceps, wet mounted using staining solution and sealed with paraffin. Photographs were taken with the 35-mm camera built into the Zeiss Axiomat. Kodak type EH-135 film was exposed for 5 s and developed at ASA 160.



system, we used contact-inhibited confluent cultures grown on glass coverslips as a model for the non-cycling state. A 'wound' scraped in a confluent monolayer releases the cells at the wound margin from contact inhibition³, thus providing both cycling and non-cycling cells on the same coverslip. To determine the maximum fluorescence difference between active and quiescent cells, we examined monolayers at a variety of dye concentrations 4 h after wounding (Fig. 1, Table 1). Table 1 shows that cycling cells were more fluorescent than non-cycling cells over a 50-fold range of dye concentration, with the intensity ratio varying from 1.5 to 5. As the maximum value of the intensity ratio was obtained with 5×10^{-7} M dye, this concentration was chosen for further studies.

The viability of cells in dye was assessed by a double staining experiment using ethidium bromide (1×10^{-6} M) as a measure of the cells' ability to exclude ethidium bromide; dead cells concentrate this dye within their nuclei, while viable cells remain unstained⁷. The cells remained viable and excluded ethidium bromide for at least 2 h in the presence of 5×10^{-7} M cyanine dye, indicating that Di-O-C₇(3) itself is not toxic in these

experimental conditions. In addition, absence of ethidium bromide staining in the cells at the wound margins indicates that cells brightly stained with cyanine dye were not damaged by the wounding process. Thus, the intensity differences shown in Table 1 are not the result of differences in viability of the cell populations.

Intensity differences were also observed in our second model system, using fibroblasts grown to different cell densities. Non-confluent and confluent cells on different coverslips were stained and measured using the previously outlined protocol. As Table 2 demonstrates, the non-confluent (cycling) cells stain more intensely than the confluent (non-cycling) cells.

Table 2 also summarizes the results obtained in the third model system, in which non-confluent cells were incubated for various periods of time in non-permissive medium (0.5% fetal calf serum), in which they become essentially non-cycling cells⁸, and measurements were made as previously described. After 6 days in this media, the cell fluorescence approached that of confluent contact-inhibited cells, although the monolayer was still subconfluent. Visual inspection of cells in low-serum medium did suggest a low level of mitotic activity.

To determine the time required for cells at the edge of a wound in a confluent layer to become significantly brighter than cells in the interior, six confluent monolayers were simultaneously wounded by scraping, returned to media and at various times later stained with 5×10^{-7} M dye for 15 min and then assayed for fluorescence intensity. The results of this experiment (Table 3) may be summarized as follows. (1) Fluorescence changes in the cells at the wound margin were noted as soon as measurements could be made, that is, within 18 min. (2) Maximal fluorescence intensity of the bright cells occurred within 18 min and a constant fluorescence ratio of wounded to confluent cells was then maintained for at least 1 h. (3) The per cent of bright cells at the wound margin increases with time and

Table 2 Effect of growth medium on fluorescence intensity in cycling and non-cycling fibroblasts

Cell density	Growth medium*	Fluorescence intensity†
Confluent monolayer	Permissive	1.00 ± 0.066
Nonconfluent monolayer	Permissive	2.89 ± 0.18
Nonconfluent monolayer	Non-permissive, 52 h	1.71 ± 0.16
Nonconfluent monolayer	Non-permissive, 6 days	1.15 ± 0.084

* Permissive = Dulbecco's modified Eagle's medium (DMEM) + 10% fetal calf serum; non-permissive = DMEM + 0.5% fetal calf serum.

† Means ± s.e. all expressed relative to the fluorescence intensity of the confluent monolayer in permissive medium measured after exposure to 6×10^{-7} M Di-O-C₇(3) as described in Fig. 1 legend.

Table 3 Changes in fluorescence of cycling and non-cycling fibroblasts with time after wounding

Time after wounding (min)	Wound margin	Confluent region	Ratio	Background
3	272 ± 25.3	55.1 ± 4.2	4.94 ± 0.39	1.22 ± 0.06
30	197 ± 15.0	42.5 ± 4.3	4.64 ± 0.59	3.48 ± 0.26
60	204 ± 16.1	46.1 ± 3.6	4.42 ± 0.49	3.35 ± 0.32

The cells were stained with 5×10^{-7} M Di-O-C₇(3) as described in the text at the indicated time after wounding. Staining required a 15-min incubation in dye at room temperature in the dark. Parameters measured were as described in Table 1 legend.

bright cells are eventually noted several cell layers into the confluent monolayer. Mitotic figures and rounded cells are observed at later times.

It is interesting to speculate on the mechanism by which the dye reflects changes in the cycling state of the cells. One possibility is that the dye stains cell cytoplasm relatively non-differentially. Observed intensity differences would then be merely the result of differences in cell size and shape between the two cell populations. Two observations oppose this interpretation. First, microscopic analysis suggests that the size and shape differences between bright cells at the wound margin and those in the confluent monolayer are minimal, especially at early time post-wounding. Second, the intensity ratio of bright to dim cells is concentration dependent (Table 1). If the fluorescent intensity difference merely reflected differences in cell size, the ratio of intensities should remain fairly constant, independent of dye concentration.

A possible mechanism for the observed fluorescence difference is that the dye stains some cytoplasmic structure. Figure 1 shows that the dye does not stain nuclear structures and is localized in discrete filamentous structures within the cytoplasm. Comparison of the morphology of staining in Fig. 1 with that obtained by Johnson *et al.*⁹ using rhodamine 123 suggests that these structures might be mitochondria. Cyanine dyes have been used to stain isolated mitochondria and seem to respond to membrane potential in such systems¹⁰. The timed wound experiments (Table 3) indicate that detection of cycling cells can be accomplished considerably earlier than is possible with methods dependent on DNA synthesis. These results also suggest that some cell parameter—perhaps the cellular or mitochondrial membrane potential—changes very soon after the cells are triggered to begin cycling.

Finally, the intensity differences reported here are large enough to allow good discrimination even in cell populations which have fairly broad fluorescence distributions. Such differences should also make the detection of even relatively small cell populations possible using quantitative techniques for single-cell analysis and sorting¹¹. Di-O-C₇(3) staining therefore promises to provide a fast, easy and sensitive method for distinguishing cycling cells from quiescent cells.

This work was supported by NIH grant GM-23088 (P. K. H., J. E. G. and K. A. M.) and a summer fellowship from the Janet E. Bently Fund of the University of Rochester Cancer Center (R.L.C.).

Received 3 October 1980; accepted 29 January 1981.

- Swartzendruber, D. E. *Expl Cell Res.* **109**, 439–443 (1977).
- Darzynkiewicz, Z. in *Flow Cytometry and Sorting* (eds Melamed, M., Mullaney, P. & Mendelsohn, M.) 285–317 (Wiley, New York, 1979).
- Cone, C. D. *Jr Ann. N.Y. Acad. Sci.* **238**, 420–435 (1975).
- Bard, J. & Wright, M. O. *J. cell. Physiol.* **84**, 111–120 (1974).
- Sims, Waggoner, A. S., Wang, C. H. & Hoffman, J. F. *Biochemistry* **13**, 3315–3330 (1974).
- Waggoner, A. S. *A. Rev. Biophys. Bioengng* **8**, 47–68 (1979).
- Horan, P. K. & Kappler, J. W. *J. immun. Meth.* **18**, 309–316 (1977).
- Pardee, A. B., Jimenez de Asúa, L. & Rozengurt, E. in *Control of Proliferation in Animal Cells* (eds Clarkson, B. & Baserga, R.) 547–562 (Cold Spring Harbor Laboratory, New York, 1974).
- Johnson, L. V., Walsh, M. L. & Chen, L. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 990–994 (1980).
- Kinnally, K. W., Tedeschi, H. & Maloff, B. L. *Biochemistry* **17**, 3419–3428 (1978).
- Horan, P. K. & Wheelless, L. L. *Science* **198**, 149–157 (1977).

Localization of pacemaking activity in early embryonic heart monitored using voltage-sensitive dye

Kohtaro Kamino, Akihiko Hirota & Shiroh Fujii

Department of Physiology, Tokyo Medical and Dental University, School of Medicine, Bunkyo-ku, Tokyo 113, Japan

Early in cardiogenesis, heart primordia are brought together at the midline and fuse with each other progressively caudally—this results in the formation of the primitive tubular heart which begins beating spontaneously at the middle period of the 9-somite developmental stage in the chick embryo¹. However, in these very early stages of development, the myocardial cells are small and technically difficult to impale with microelectrodes; thus electrophysiological studies on the very early embryonic heart are rare^{2,3}. Recently, potential sensitive dye-related absorption signals have provided a new method for monitoring spontaneous action potential activity in the early embryonic heart^{4–7}. This technique is based on the observation that changes in potential across membrane(s) stained with certain voltage-sensitive dyes are accompanied by changes in their optical properties (absorption, fluorescence, and/or birefringence)^{8–10}. Using absorption signals, we have already demonstrated in embryonic pre-beating chick heart in the 7–8-somite stages, the occurrence of action potential activity^{4,5}, development of pacemaker potential^{6,7} and cardiac rhythm generation⁷. With this method, originally introduced to record neuronal activity in invertebrate ganglia^{11,12}, many cells or portions of the preparation can be monitored simultaneously. Accordingly we have expanded the optical recording apparatus to monitor simultaneously spontaneous action potentials from five portions of an early embryonic heart, and report here experiments carried out on the embryonic hearts of chicks (white Leghorn) at the 7–11-somite developmental stages, corresponding to 25–35 h of incubation. The hearts attached to the embryo were stained with a merocyanine-rhodanine dye (NK2761) as a potentiometric probe. This dye is an analogue of Dye XVII⁹ or Dye XXIII¹⁰.

Figure 1 shows simultaneous recordings of spontaneous optical signals from an 8-somite embryonic heart. The circles overlaid on the view of the embryonic heart in this figure are drawn to scale, indicating the field of optical recording and the relative positions of the light guides. Traces 1, 2 and 4 were obtained from the ventricular region, and traces 3 and 5 from the unfused atrial levels. In the preparation shown in Fig. 1b, both size and shape of each spike in individual traces are approximately uniform, which suggests that the action potentials of active cells in the field of optical recording were well synchronized. However, as shown in traces at the higher sweep speed, there were obvious differences in time course of action potentials from the five different regions; the action potentials in the five different regions were synchronized and there was a small delay between spike 5 and that in the other regions. It is thus likely that the delay corresponds to spreading time of the action potentials. At the 7–8-somite developmental stages, the action potential was usually first detected in the left atrial primordium, but sometimes in the right primordium.

Figure 2 shows experiments demonstrating the spatial distribution of spontaneous action potential activity in the 9–11-somite embryonic hearts. In the preparation shown in Fig. 2a, the action potential first appeared in the left atrial portion (trace 5), and action potentials in other portions (traces 1–4) followed those in the left atrial portion after a delay of 40–150 ms. Furthermore, the light guides were positioned close together in a series on the left portion of a 9-somite embryonic heart. In this experiment, the action potential probably occurs first in the left

atrial level and then spreads progressively in rostral and caudal directions. As shown in Fig. 2b, an action potential also appeared first in the left atrial region in the heart of the 10-somite embryo. Thus, it is reasonable to assume that although both the right and left atria have intrinsic pacemaking abilities, the cardiac rhythm as a whole is governed by the left atrium as the embryo develops from the 7- to the 10-somite stage.

From experiments on the pulsation rate, Patten¹³ suggested that the cardiac pacemaker at the developmental stage of initiation of heartbeat is the atrium. Using microsurgical techniques, DeHaan¹⁴ showed that the earliest pacemaker was on the left

side, and that it shifted to the sinoatrial region only later. Moreover, Van Mierop² was able to record the earliest action potentials in the left atrial region of the 11-somite beating heart, and in the left atrial primordium in the 8-somite embryo, before contractile activity had begun. Indeed, his transmembrane-potential records are remarkably similar to our optical recordings (Fig. 1b). As development proceeded to the 11-somite stage (Fig. 2c), the portion in which the action potential first occurred shifted caudally along the atrium to the sinus primordium.

In embryonic chick hearts past the 13-somite stage, pacemaker potentials were detected electrophysiologically using microelectrodes^{15,16}. We have optically demonstrated differentiation and development of the pacemaker potential during the 8–9-somite stages^{6,7}. At the 8-somite stage, absorption signals resembling the pacemaker potential are widely distributed in the embryonic hearts but by the beginning of the 9-somite stage these signals are localized at the atrial level¹⁷.

In the present experiment, the cardiac-type signals were obtained from the ventricular region and the pacemaker-type ones from the atrial region. The localization of the latter is consistent with the results obtained from experiments on the time course of action potentials, and shows clearly that the pacemaking area is situated in the atrium region at the 9-somite stage.

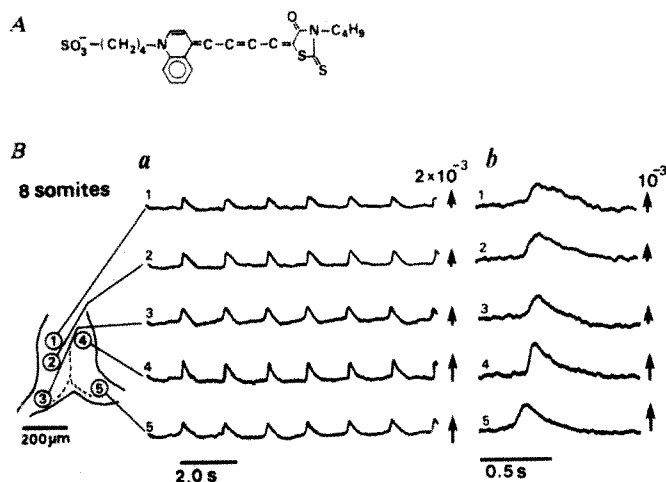


Fig. 1 A, Chemical structure of a merocyanine-rhodanine dye (NK2761, produced by Nippon-Kankoh-Shikiso Kenkyusho Co.) used as a potentiometric probe in our experiments. The cation associated with the dye is sodium. Voltage clamp experiments in squid giant axons gave good evidence that the absorption signals of merocyanine-rhodanine dyes were linearly related to membrane potential change, and that the optical response followed the voltage change with a time constant of $<20 \mu\text{s}$ (refs 9, 10 and L. B. Cohen, personal communication). B, Spontaneous absorption signals detected simultaneously from five regions in an 8-somite embryonic heart. After the splanchnopleure was carefully peeled off, the isolated embryos were incubated for 15 min in bathing solution containing 0.05 mg ml^{-1} of the dye. After incubation, the dye solution was washed out. The dye was bound to the epi-myocardium cells. The solution used to bathe the embryos was 138 mM NaCl, 5.4 mM KCl, 1.8 mM CaCl₂, 0.5 mM MgCl₂, and 10 mM Tris-HCl buffer (pH 7.2), and was equilibrated with air and allowed to warm to $\sim 37^\circ\text{C}$. This solution was used in most experiments, except those shown in Figs. 2B and C. To measure the absorption of light by the area of interest in the dyed embryonic heart, a $\times 50$ enlarged image of the heart was formed with a $\times 20$, 0.46 n.a. (numerical aperture) objective and a $\times 2.5$ photographic eyepiece, using an Olympus-Vanox microscope (AHB-LB-1, Olympus Optical Co.). Five light guides were used to carry the light from the image of the embryonic heart to each photodiode (Bell and Howell, 509-10). The tips of these guides (4.0 mm in diameter) were positioned over the image of the embryonic heart stained with the dye. The measuring beam was produced by a JC 24V/300W halogen-tungsten-filament lamp (Kondo Sylvania) powered by a current regulated power supply (PAD 35-20L, Kikusui Electronics); the incident light was passed through an interference filter with a centre wavelength of 700 nm and a width at half-height of 11 nm (1F-S, Vacuum Optical Japan). The signals in B were eliminated by illumination with white light, or at 620 nm where the action spectrum of the absorption signal of the merocyanine-rhodanine dyes is zero^{6,7}. Thus, the optical signals shown resemble the spontaneous action potentials in the embryonic heart. The empirical criteria required to distinguish between optical signals from an action potential and those corresponding to movement artefacts have been described previously⁵. In experiments where a relatively low light intensity and brief duration were used, photodynamic damage and dye bleaching were not serious. The fractional absorption change ($\Delta A/A$) is related to the fractional change in transmitted intensity (T); evidence presented below indicates that a decrease in intensity was due to an increase in absorption. $\Delta A/A$ is equal to $-\Delta T/(T_{\text{before}} - T_{\text{after staining}})$, ($\Delta T/T_{\text{after}} = \Delta I/I$). Here, the transmittance calibration bar is indicated. The traces were detected in a single sweep at 37.6°C , and the oscilloscope (Tektronix 5113 Dual Beam Storage Oscilloscope) was set to give a coupling time constant of 1.5 s. The outputs were finally filtered by a simple RC low-pass filter (time constant $\sim 10 \text{ ms}$). Positions 1, 2 and 4 are from the ventricular region, and positions 3 and 5 on the atrial unfused primordia. The field of optical recording encircled in the figure is estimated as 0.005 mm^2 . Traces b were obtained at the higher sweep speed. The embryonic heart is shown in ventral view.

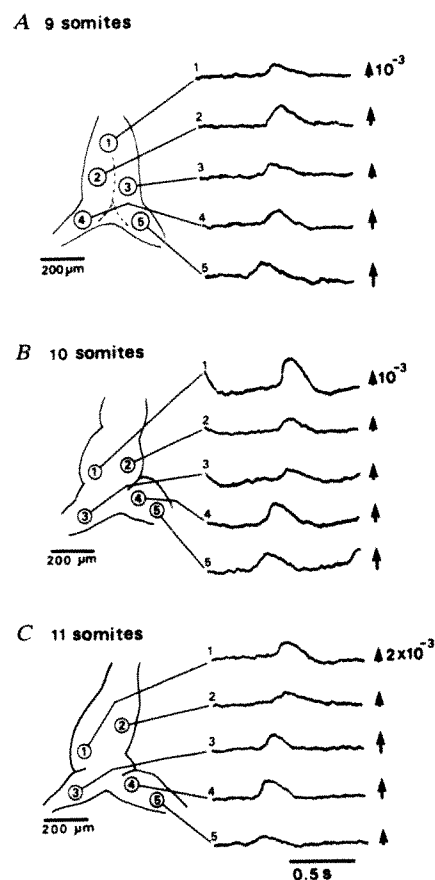


Fig. 2 Simultaneous absorption measurements in five different portions in 9-(a), 10-(b) and 11-(c) somite embryonic chick hearts. Light-guides 1, 2 and 3 were positioned over the image of the ventricular region, and 4 and 5 over the image of the atrial unfused regions in the experiment shown in A. This preparation was obtained just before initiation of the heartbeat⁵. In B and C, traces 1 and 2 were obtained from the ventricular region, 3 and 4 from the atrium, and the position of trace 5 corresponds to sinus primordia. At these stages, the hearts are bent to the right and are pulsating. To reduce movement artifact due to pulsation, we used a hypertonic solution made by increasing NaCl concentration up to 276 mM. In this solution, contraction was reduced. However, such hearts are electrically excitable⁶, as is the case in striate muscle¹⁸. Thus, it is not surprising that no signal was observed by illumination with white light or at 620 nm. It seems likely that differences in the signal size among the different positions are due to tissue thickness and density of the active membrane⁶. Experiments were carried out at 37.3°C (A), 38.3°C (B) and 36.7°C (C).

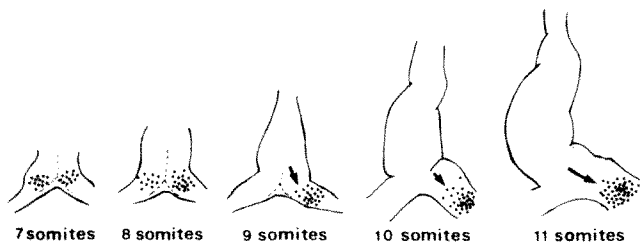


Fig. 3 Maps summarizing the location of the pacemaking area in the embryonic chick hearts throughout early phases of development, from the 7- to the 11-somite stage. Depth of dotting represents the degree of priority of pacemaking. The arrows indicate the direction of translocation of pacemaking area.

On the basis of these results, we constructed a rough developmental map of regional localization of pacemaking areas during the early phases of cardiogenesis (Fig. 3). The possibility exists that pacemaker cells are situated originally in the cardiac primordia just before and after they fuse in the midline. It seems that as the primitive tubular heart is formed, the pacemaking area which possesses the highest intrinsic rhythmicity, is localized mainly in the left atrial primordium and subsequently in the sinus primordium. With further development of the heart, the

pacemaking cells of the sinus primordium probably become organized into nodes and bundles while maintaining their original relationship with the surrounding musculature.

We thank Drs L. B. Cohen and H. Irisawa for reading the manuscript and for comments and suggestions, and M. Ohara for useful advice. We also thank Nippon Kankoh-Shikiso Kenkyusho Co. for synthesizing the merocyanine-rhodanine dye. The work was supported by grants 457021, 422007 and 387039 from the Ministry of Education, Science and Culture of Japan.

Received 4 September 1980; accepted 27 January 1981.

1. Patten, B. M. *Physiol. Rev.* **29**, 31–47 (1949).
2. Van Mierop, L. H. S. *Am. J. Physiol.* **212**, 407–415 (1967).
3. Bernard, C. in *Developmental and Physiological Correlates of Cardiac Muscle* (eds Lieberman, M. & Sano, T.) 169–184 (Raven, New York, 1975).
4. Hirota, A., Fujii, S. & Kamino, K. *Jap. J. Physiol.* **29**, 635–639 (1979).
5. Fujii, S., Hirota, A. & Kamino, K. *J. Physiol., Lond.* **304**, 503–518 (1980).
6. Fujii, S., Hirota, A. & Kamino, K. *J. Physiol., Lond.* **311**, 147–160 (1981).
7. Fujii, S., Hirota, A. & Kamino, K. *J. Physiol., Lond.* **312** (in the press).
8. Cohen, L. B. *et al. J. Membrane Biol.* **19**, 1–36 (1974).
9. Ross, W. N. *et al. J. Membrane Biol.* **33**, 141–183 (1977).
10. Gupta, R. K. *et al. J. Membrane Biol.* (in the press).
11. Salzberg, B. M., Grinvald, A., Cohen, H. V., Davila, H. V. & Ross, W. N. *J. Neurophysiol.* **40**, 1281–1291 (1977).
12. Grinvald, A., Cohen, L. B., Leshner, S. & Boyle, M. B. *J. Neurophysiol.* (in the press).
13. Patten, B. M. *Univ. Mich. med. Bull.* **22**, 1–21 (1956).
14. DeHaan, R. L. *Dev. Biol.* **1**, 586–602 (1959).
15. Meda, E. & Ferronie, A. *Experientia* **15**, 427–428 (1959).
16. Lieberman, M. & Paes DeCarvalho, A. *J. gen. Physiol.* **49**, 351–363 (1965).
17. Fujii, S., Hirota, A. & Kamino, K. *J. Physiol., Lond.* (submitted).
18. Hodgkin, A. L. & Horowitz, P. *J. Physiol., Lond.* **136**, 17P–18P (1957).

Direct communication between axons and sheath glial cells in crayfish

Camillo Peracchia

Department of Physiology, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642, USA

Evidence for the transfer of molecules as large as proteins from glial or Schwann cells to axons has frequently been reported^{1–8}. This phenomenon is well supported in crayfish^{3,4,8} and squid^{5–7} but there is some doubt about its existence in vertebrates⁹. In invertebrates most of the data support the transfer from glial cells to axons, but little is known about movement of molecules in the opposite direction. However, Grossfeld *et al.*¹⁰ recently reported data on the passage of a fluorescent dye (Lucifer yellow CH, molecular weight (MW) 476) from a crayfish median giant axon to all glial cells of the adaxonal layer. Little is known of the mechanism by which molecules are exchanged between axons and glial cells, but the bidirectionality of the phenomenon in crayfish axons suggests that channels or pores for direct intercellular communication may exist. In most cells, direct communication is provided by gap junctions¹¹ but these structures have never been seen between axons and glial cells of crayfish^{12–14}. Indeed, regions of plasma membrane specialization have been reported¹³, but they are structurally different from gap junctions. I describe here pore structures, observed between 51 axons and glial cells of six crayfish spinal cords, which may represent the means of communication between these cells.

In crayfish, *Procambarus clarkii*, anaesthetized by cooling, the sternal artery was cannulated using polyethylene tubing and perfused with 50–100 USP units of heparin diluted in 10 ml of Van Harreveld's salt solution¹⁵, followed by 40–60 ml of 3–6% glutaraldehyde solution buffered to pH 7.4 with 0.1 M (Na, K) phosphate at room temperature¹⁴. The chain of abdominal ganglia was removed from the animal, washed briefly in 0.2 M phosphate and post-fixed for 2 h in 2% osmium tetroxide buffered with phosphate to pH 7.4 at room temperature. Dehydration was in graded alcohol and embedding in Epon.

Electron microscope examination of thin sections through axons of the abdominal ganglia often show regions where axonal

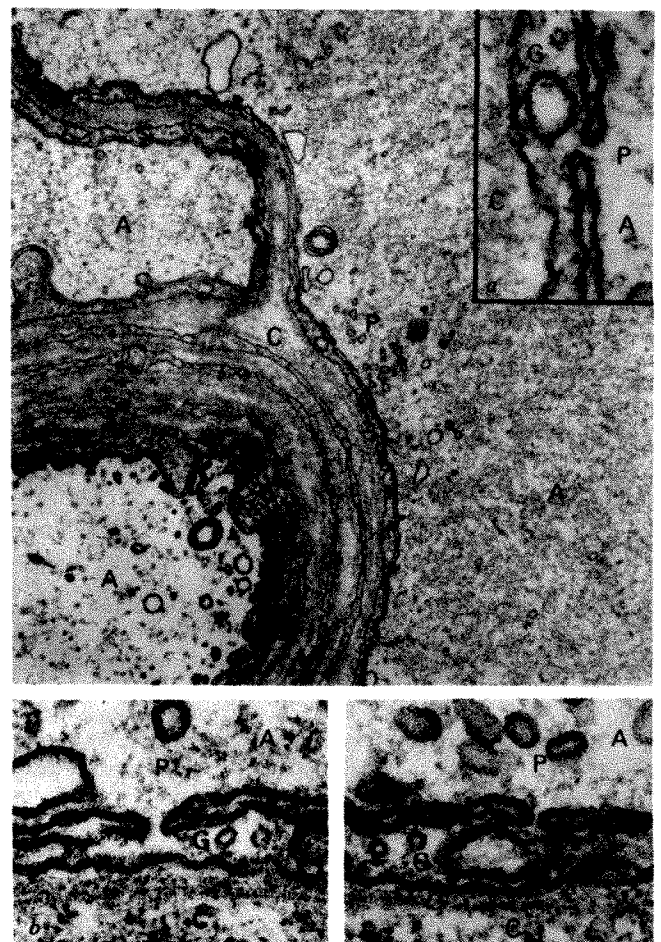


Fig. 1 Thin section through nerve fibre of a crayfish abdominal ganglion. Axonal and glial cell plasma membranes fuse with each other to form a well defined pore (P). In the region of the pore, tubules of endoplasmic reticulum are aggregated. Inset a is an enlargement of the pore shown in the low power micrograph. Insets b and c show similar pores in other axons, A, axon; G, glial cell; C, connective tissue. $\times 18,000$; insets $\times 77,000$.

and glial cell plasma membranes fuse with each other to create well defined pores between axoplasm and glial cell cytoplasm (Fig. 1). Several such pores are usually grouped in the same region. In these regions there are always aggregates of endoplasmic reticulum tubules and other membrane structures (Fig. 1). The pores generally have 15–20 nm diameters but larger pores are also seen (Fig. 1, insets).

Pores of this size are expected to allow wide communication between axoplasm and glial cell cytoplasm. It seems unlikely that such structures are artefacts of the preparative procedure, as the membranes and other cellular structures of these perfused samples were always extremely well preserved. Furthermore, if the pores were artefacts, one would expect to see them anywhere and not just at regions of endoplasmic reticulum accumulation. Indeed the existence of pores between axons and Schwann cells of invertebrates seems reasonable from a theoretical viewpoint, because only passageways of this size could allow exchange of molecules such as RNA³ and proteins as heavy as 200,000 MW (refs 5, 6).

This work was supported by NIH grants GM 20113 and NB 07107.

Received 27 October 1980; accepted 30 January 1981.

1. Singer, M. & Salpeter, M. J. *Morph.* **120**, 281–316 (1966).
2. Singer, M. & Green, M. R. J. *Morph.* **124**, 321–344 (1968).
3. Andersson, E., Edstrom, A. & Jarlstedt, J. *Acta physiol. scand.* **78**, 491–502 (1970).
4. Bittner, G. *Am. Zool.* **13**, 379–408 (1973).
5. Lasek, R. J., Gainer, H. & Przybylski, R. J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1188–1192 (1974).
6. Lasek, R. J., Gainer, H. & Barker, J. L. J. *Cell Biol.* **74**, 501–523 (1977).
7. Gainer, H., Tasaki, I. & Lasek, R. J. *J. Cell Biol.* **74**, 524–530 (1977).
8. Meyer, M. R. & Bittner, G. D. *Brain Res.* **143**, 213–232 (1978).
9. Gambetti, P., Autilio-Gambetti, L. & Peck, K. *Brain Res.* **200**, 59–68 (1980).
10. Grossfeld, R. M., Bittner, G. D., Ballinger, M. L. & Viancour, T. A. *Soc. Neurosci. Abstr.* **6**, 385 (1980).
11. Peracchia, C. *Int. Rev. Cytol.* **66**, 81–146 (1980).
12. Peracchia, C. *J. Cell Biol.* **44**, 125–133 (1970).
13. Peracchia, C. *J. Cell Biol.* **61**, 107–122 (1974).
14. Peracchia, C. & Robertson, J. D. *J. Cell Biol.* **53**, 223–239 (1971).
15. Van Harreveld, A. *Proc. Soc. exp. Biol. Med.* **34**, 428–432 (1936).

Potassium channels in the nodal membrane of rat myelinated fibres

Ofer Binah & Yoram Palti

Department of Physiology and Biophysics, Technion—Faculty of Medicine, Haifa, Israel

Following some preliminary reports^{1,2}, mammalian fibres from rabbit³ and rat^{4,5} have recently been successfully studied in detail by means of the voltage clamp. The early transient or sodium conductance system was found to be similar to that in frog and squid axons. However, the delayed conductance or potassium currents were found to be negligible^{4,6}. Only after chemical and osmotic manipulations, which were said to expose channels buried under the myelin, did Chiu and Ritchie⁷ find delayed currents in rabbit fibres. If confirmed, this would mean that the membrane conductance system of mammalian fibres is so different from that of invertebrate and amphibian axon models as to make the data base gathered from amphibian myelinated fibres (frog and toad) and invertebrate giant axons (squid and myxocolla) irrelevant to human and other mammalian fibres. However, we show here that it is possible to find in the normal nodal membrane of rat myelinated fibres potassium currents that flow through channels which are similar in many respects to those found in the frog node or squid axons.

Single myelinated motor fibres isolated from the sciatic nerve of rats were voltage clamped following the procedure of, and using the basic apparatus described by Nonner⁸. The voltage

clamp command potentials as well as the membrane current sampling rates were under digital computer (NOVA 800) control⁹. Holding potential (V_H) was adjusted to give a 70% sodium conductance inactivation. Membrane potential (V_M), which was defined at the end of each experiment by destroying the nodal membrane, averaged -78 ± 5 (s.d.) mV in 13 preparations. All currents were corrected for leakage (I_L) by computing the leak conductance from the inward currents elicited by membrane hyperpolarization. It was assumed that the leakage current–voltage relationship is linear and that $I_L = 0$ at the resting potential. The node was externally perfused with mammalian Ringer's solution. Potassium currents were measured in the presence of 300 nM tetrodotoxin (TTX). The composition of the internal solution was controlled by diffusion through the fibre's ends which were cut 400–700 μm away from the node¹⁰. Currents were converted to densities (mA cm^{-2}) assuming the nodal area to be 50 μm^2 (ref. 9).

Current tracings obtained in TTX Ringer at normal holding potentials, V_H , are shown in the insert of Fig. 1a. The tracings show a large instantaneous outward current jump (I_{inst}) followed by a delayed current component. I_{inst} was measured by extrapolating (on semilogarithmic scale) the initial position of the current⁷ tracing, beyond the capacity transient, to zero time.

The instantaneous current shows a linear $I_{\text{inst}}-V$ relationship, while the steady-state current, $I_{\text{ss}}-V$ curve shows some rectification and is very similar to that of the frog node (Fig. 2). Note that following the relatively rapid increase of the current in the first 3 ms there is a steady slow increase in current. This additional slow activation can be best demonstrated by the change of slope of the $\log(I_{\infty} - I_L)$ versus t curve as seen in the insert of Fig. 3b. In the example given, the initial time constant, τ_n , is 0.53 ms (this value is obtained after 'peeling off' the second time constant) and the slower time constant, τ_s , is 1.36 ms. The corresponding single τ at $V_H = -71$ mV is 0.56 ms, that is, τ_n is not altered by depolarizing V_H . Shifting V_H in the hyperpolarizing direction ($V_H = -100$ mV) has very little effect on the instantaneous current and produces a 15% increase in the delayed currents (open symbols in Fig. 1a). However, a relatively small (9 mV) shift of V_H in the depolarizing direction has a marked effect on the current tracings (Fig. 1b): they become square and are thus similar to those described by Chiu *et al.*⁶ and Brismar⁴. The squaring of the current traces, which was completely reversible, results mainly from an increase of the I_{inst} values to approach those of the steady state (I_{ss}). I_{ss} remains unchanged for depolarizations up to +10 mV, beyond which the currents generated at $V_H = -71$ mV are reduced by 10–15%. Note, however, that most of this reduction is due to the disappearance of the slow activation (see Fig. 1b). This disappearance or inactivation of the slow process obviously also contributes to the 'squaring off' of the currents.

Figure 2 compares the steady-state $I_{\text{ss}}-V$ relationship in the frog and rat. The shape of the curves obtained by means of the same voltage-clamp setup and procedure is clearly very similar, although the frog curve shows more rectification.

Moreover, in both cases the zero-current voltage cross-over point is shifted from around rest in normal Ringer's towards zero in high external $[\text{K}^+]$, indicating that the current is carried mainly by K^+ ions. Note the difference between the effect of high $[\text{K}]_o$ on the two preparations: it increases the potassium conductance, g_K , by about 50% in the frog and reduces g_K by 30% in the rat. The assumption that K^+ ions are the main charge carriers in the delayed channels is strengthened by the fact that the axoplasm in our fibres was in equilibrium with isotonic KCl solution. Therefore, the only membrane permeable ions in the axoplasm must be K^+ and Cl^- (ref. 10). As the chloride concentration remained unchanged, the large shift of the reversal potential indicates that the chloride permeability is relatively low. This low chloride permeability which is also typical of the frog node¹¹, indicates that the outward currents we are dealing with must be mainly carried by K^+ ions.

As reported by Chiu⁶ and Brismar⁴, we also find that external tetraethylammonium (TEA) only partly attenuates I_{ss} even at

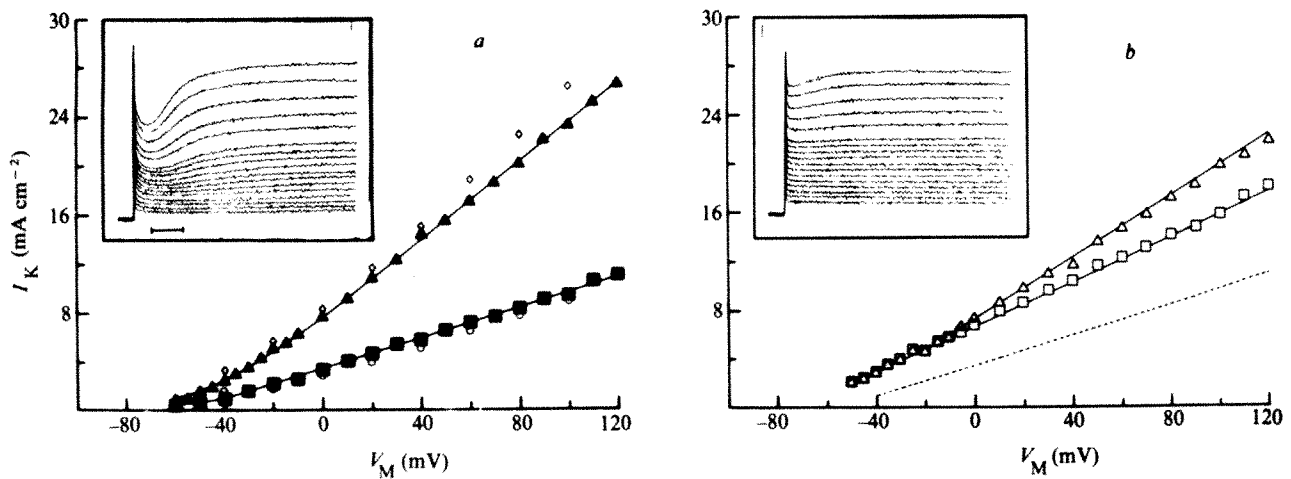


Fig. 1 Steady-state and instantaneous current-voltage relationships at various holding potentials (V_H). All currents corrected for leakage. Test solution: Ringer TTX. Temperature 20 °C. *a*, $\blacktriangle$, Steady-state currents at $V_H = -80$ mV. $\blacksquare$, Instantaneous currents at $V_H = -80$ mV. $\diamond$, Steady-state currents at $V_H = -100$ mV. $\circ$, Instantaneous currents at $V_H = -100$ mV. Inset illustrates the recorded membrane currents as a function of time for different depolarizing pulses. The potential increment between the pulses was 5 mV from 15 to 60 mV and 10 mV from 60 to 120 mV. Calibration: horizontal bar = 0.8 ms. *b*, $\triangle$, Steady-state currents at $V_H = -71$ mV. $\square$, Instantaneous currents at $V_H = -71$ mV. The interrupted line gives the instantaneous currents generated by the same fibre at $V_H = -80$ mV. Inset illustrates the recorded membrane currents as a function of time for different depolarizing pulses. The pulse amplitudes are the same at those of *a*, as is the time scale.

relatively high concentrations (10 mM TEA attenuates I_{ss} by only 30%). I_{inst} was practically unaffected by TEA.

Figure 3a illustrates the voltage dependency of the normalized delayed channel steady-state conductance parameter, n_{∞}^x , which was computed using the following relationship:

$$n_{\infty}^x = I_{ss} / \bar{g}_K (V_M - V_K) \quad (1)$$

where $\bar{g}_K$ is the maximal membrane potassium conductance, V_K is the delayed channel reversal potential and x is the power of n in the Hodgkin-Huxley axon model. It is seen that near the resting potential ($V_M = -60$ mV) the potassium conductance, as determined by n_{∞}^x , is far from zero for both $x = 4$ (circles) and $x = 1$ (triangles). This is in contrast to the frog node or squid, where for the experimentally determined $x = 4$, the potassium conductance is negligible (interrupted line). The relatively high potassium conductance at the resting potential is obviously consistent with the relatively large initial current jump 'squaring off' obtained here and by others⁴ on membrane depolarization. It is also consistent with the initial current decay observed on membrane hyperpolarization by Chiu *et al.*⁶, and confirmed in the present work. The rate of this decay conforms with the projected values of τ_n at these potentials.

On the basis of the above, we suspect that the conclusion of Chiu *et al.* that (in the rabbit) this decaying current is capacitive (I_C) may be wrong. Moreover, because $I_C = C_M [dV_M/dt]$ (ref. 12), as $C_M \neq 0$ one must conclude that in their experiment $dV_M/dt \neq 0$, that is, the membrane was not fully voltage clamped throughout the first 0.5–1 ms.

We cannot attribute the almost completely square currents recorded by Chiu⁶ in the rabbit, and Brismar⁴ in the rat, to the shifted n_{∞} curve alone. However, this shift together with a somewhat depolarized membrane, as suggested by Chiu *et al.* themselves, may explain the results. Within the above framework, we may interpret the appearance of potassium currents following the rather drastic chemical and osmotic treatments in the rabbit⁶ as being due to a shift in the relative holding potential and perhaps also to a modification of the slow potassium activation process.

To rule out the possibility that in our experiments the myelin was lifted from the axolemma at the internode, we measured the nodal electrical capacity using the triangular wave voltage-clamp method¹². The nodal capacitance was found to be about 2.7 pF, compatible with that of the frog node¹³. The capacity did not change by more than 10% during the V_H changes which

induced the potassium conductance changes. These small C_M changes in the rat are definitely incompatible with the apparent 10-fold C_M change reported in the rabbit by Chiu *et al.*⁷ on the basis of the transient inward current obtained during membrane hyperpolarization.

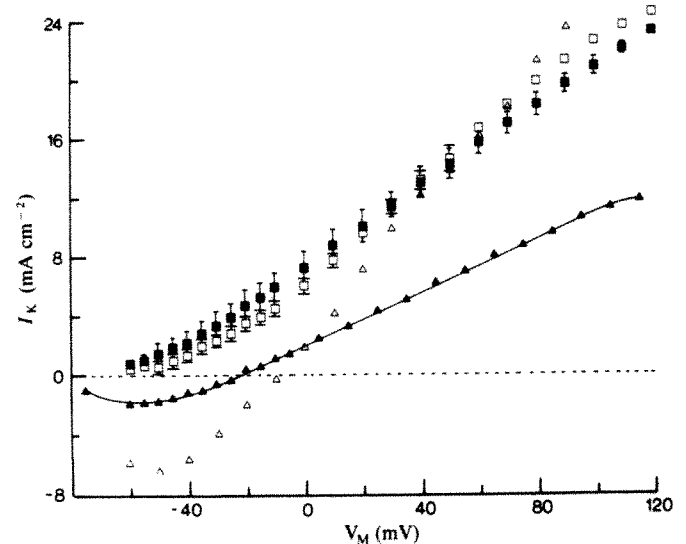
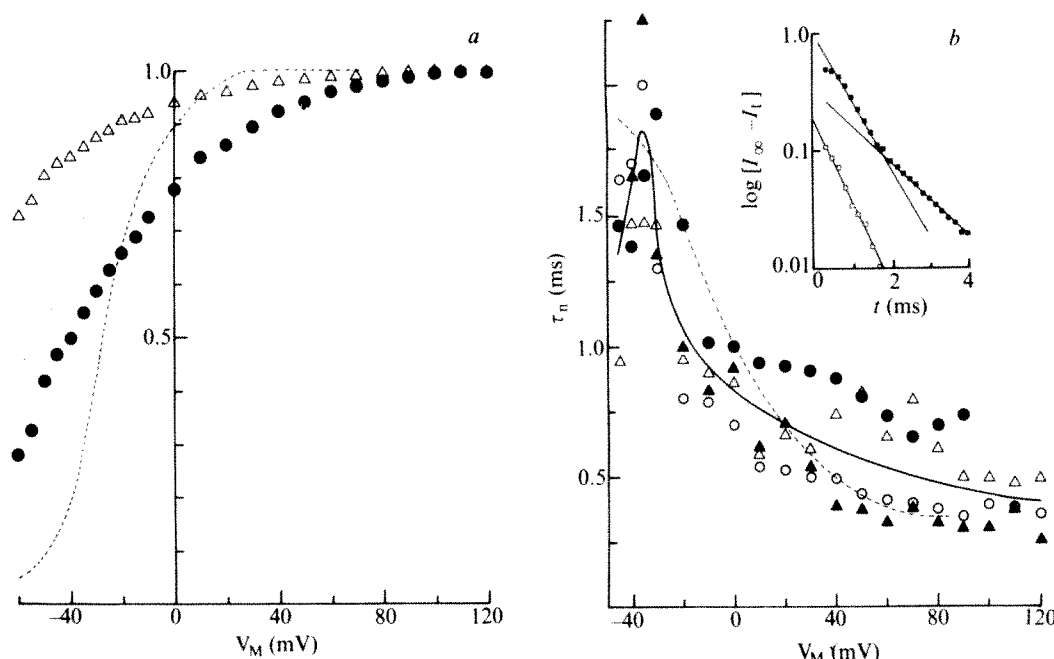


Fig. 2 Potassium current-voltage relationships measured from frog and rat nodes by means of the same voltage clamp. All currents corrected for leakage. Temperature 20 °C. $\blacksquare$, Potassium currents from rat node. Each symbol represents an average current value from six fibres $\pm$ s.d. To eliminate scatter of data due to 'g' effect (through, for example, variance in the number of channels) all currents of each fibre, were first normalized by multiplying each current value by the ratio of the current value at $V = 120$ mV to the average current value for all fibres at this potential. $\square$, Potassium currents from the frog node (*Rana ridibunda*). Each symbol represents an average current value from four fibres $\pm$ s.d. Current normalization was carried out as above. $\blacktriangle$, Potassium currents measured from a rat node when the external Ringer solution was replaced by an isotonic KCl solution. $\triangle$, Potassium currents from a frog (*Rana esculenta*) when the external potassium concentration was raised from 2.5 to 80 mM.

Fig. 3 Potassium conductance parameters in the rat node. Temperature 20°C. *a*, The normalized steady-state potassium conductance (fraction of open channels). ●, n_{∞}^x values calculated from equation (1) for $x = 4$. △, n_{∞}^x values calculated from equation (1) for $x = 1$. Interrupted line: normalized conductance or fraction of open channels in frogs ($x = 4$), after Hille¹⁴. Note that close to the resting potential, at $V = -60$ mV, the relative conductance of the rat is at least five times larger than that of the frog. *b*, The potassium conductance time constants, τ_n , versus membrane potential. The symbols represent four rat fibres. Solid line, the average τ_n . Interrupted line, τ_n values in frog taken after Hille¹⁴. Inset: the $\log [I_{\infty} - I_t]$ versus time (t) relationships for potassium currents at $V_H = -80$ mV (filled symbols) and $V_H = -71$ mV (open symbols). Note the double exponential at $V_H = -80$ mV, the disappearance of the second exponential when $V_H = -71$ mV and the similarity between the first exponential at -80 mV and the single exponential at -71 mV. Note also in Fig. 1*b* that beyond 2 ms the current traces for $V_H = -71$ mV become completely flat so that we could not continue the corresponding trace beyond this time.



Beyond the initial delay, the potassium currents show double exponential kinetics when plotted on a semilogarithmic scale (Fig. 3*b* inset). From the slopes of the initial exponential relationships we computed the potassium conductance time constant, τ_n . Figure 3*b* shows the voltage dependence of τ_n from four fibres. The shape of the curve and the absolute τ_n values are similar to those previously measured in frog¹⁴ (broken line) and in rabbit 'internode'⁶.

Thus, when compared with amphibian fibres, the delayed channels in the rat are characterized by the following features: (1) The n_{∞} - V_M curve is shifted towards more negative potentials, that is, at negative membrane potentials the n_{∞} values are much higher than the corresponding values in the frog and squid. This fact is responsible for the much larger conductances near the resting potential and as a consequence a large initial outward current step on depolarization. Many of the special properties of the potassium system in the rat, such as the current decay on membrane hyperpolarization, are probably due to this fact. (2) The values and voltage dependence of the initial time constant of the change in potassium conductance, τ_n , are very similar to the corresponding parameters in frog. This parameter is not affected by depolarizing V_H . (3) The initial increase in I_K , obtained on membrane depolarization, is followed by a smaller and slower increase or activation. The time constant of this slower activation process is about three times larger than the initial ' τ_n ' value. (4) Prolonged small membrane depolarizations (5–10 mV) attenuate steady-state I_K by 10–15% mainly by reducing the amplitude of the slow activation to virtually zero. The time constant of the development of this inhibition is >20 ms. Prolonged membrane hyperpolarizations increase steady-state I_K by 10–20%.

On the basis of the above findings, we believe that the excitability properties (threshold, action potential, refractoriness) of mammalian fibres in normal conditions can be analysed and reconstructed on the basis of models, similar in principle but modified in some details, with regard to the potassium system, to those developed for the frog and squid.

This study was performed in partial fulfillment of the requirements of the D.Sc. degree for O.B.

Received 18 July 1980; accepted 7 January 1981.

- Horackova, M., Nonner, W. & Staempfli, R. *Proc. int. Un. physiol. Sci.* **7**, 198 (1968).
- Nonner, W. & Staempfli, R. in *Laboratory Techniques in Membrane Biophysics* (eds Passow, H. & Staempfli, R.) 171–175 (Springer, Berlin, 1969).
- Chiu, S. Y., Mrose, H. E. & Ritchie, J. M. *Nature* **279**, 327–328 (1979).
- Brismar, T. J. *Physiol., Lond.* **298**, 171–184 (1980).
- Palti, Y. & Binah, O. *Biophys. Soc. Abstr.*, 2074 (1980).
- Chiu, S. Y., Ritchie, J. M., Rogart, R. B. & Stagg, D. J. *Physiol., Lond.* **292**, 149–166 (1979).
- Chiu, S. Y. & Ritchie, J. M. *Nature* **284**, 170–171 (1980).
- Nonner, W. *Pflügers Arch. ges. Physiol.* **309**, 116–192 (1969).
- Palti, Y., Moran, N. & Staempfli, R. *Biophys. J.* (in the press).
- Palti, Y., Gold, R. & Staempfli, R. *Biophys. J.* **25**, 17–31 (1979).
- Århem, P., Frankenhaeuser, B. & Moore, L. E. *Acta physiol. scand.* **88**, 454–466 (1973).
- Palti, Y. & Adelman, W. J. *J. Membrane Biol.* **1**, 431–458 (1969).
- Hodgkin, A. L. *The Conduction of the Nervous Impulse* (Springfield, Illinois, 1964).
- Hille, B. in *Biophysics and Physiology of Excitable Membranes* (ed. Adelman, W. J.) 230–246 (Van Nostrand, New York, 1980).

Prolactin inhibits oestrogen synthesis in the ovary

Jennifer Dorrington & Robert E. Gore-Langton

Banting and Best Department of Medical Research,
University of Toronto, Toronto, Ontario, Canada M5G 1L6

In 20% of cases of secondary amenorrhoea, hyperprolactinaemia prevents ovulation by impairing normal follicular development, but little is known of the biochemical basis for this effect¹. Bromocriptine can restore follicular growth and ovulation² by inhibiting the release of prolactin from the pituitary. The suckling stimulus causes an increase in prolactin levels³, and ovarian follicles fail to develop fully, thus inducing an anovulatory state throughout lactation in many mammals. We report here experiments with cultured granulosa cells which suggest that this contraceptive action of prolactin is due to its ability to interfere with the action of follicle-stimulating hormone (FSH) on the synthesis of oestrogen.

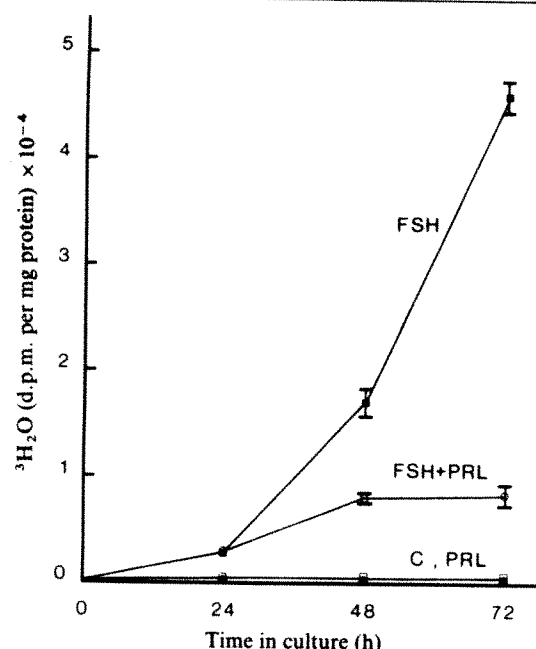


Fig. 1 Effect of FSH and prolactin on the aromatase activity of ovarian granulosa cells, assessed by the release of $^3\text{H}_2\text{O}$ from $[1\beta\text{-}^3\text{H}]$ testosterone. Granulosa cells were obtained from 25-day-old rats which had been treated daily for 3 days with 1 mg diethylstilboestrol in sesame oil. The cells were plated in Falcon 24-well tissue culture plates so that each well contained 100 μg protein and 0.5 ml modified Eagle's minimum essential medium, and were maintained at 37 °C in a water-saturated atmosphere of 95% air and 5% CO_2 . Cells were treated with 300 ng NIH-FSH-S13 ml^{-1} and/or 1 μg NIH-P-S-7 ovine prolactin ml^{-1} 2 h after plating; the medium was changed after 24 and 48 h and appropriate treatments repeated. At 24, 48 and 72 h, 0.1 μCi $[1\beta\text{-}^3\text{H}]$ testosterone, containing unlabelled testosterone to give a final concentration of 0.25 μM , was added to four replicate cultures from each treatment group (C, controls; PRL, prolactin-treated; FSH, FSH-treated; FSH+PRL, FSH and prolactin-treated). The cells were incubated for 2 h at 37 °C and the reaction stopped by placing the plates on ice. The medium was transferred to tubes and Norit A-activated charcoal (25 mg in 100 μl H_2O) was added. Incubations with charcoal for 2 h at 0 °C completely adsorbed steroids without influencing the recovery of tritiated water. Steroids adsorbed to charcoal were removed by centrifuging at 4,000 r.p.m. and 4 °C for 30 min. An aliquot (0.5 ml) of the supernatant was counted in 5 ml Aquasol.

The aromatization of the androgens, androstenedione and testosterone, to oestrogens in the ovarian granulosa cells, under the influence of FSH, involves a series of closely integrated reactions catalysed by an enzyme complex referred to as 'aromatase'⁴. To assess the aromatase activity of granulosa cells after various hormone treatments, the stereospecific release of $^3\text{H}_2\text{O}$ from $[1\beta\text{-}^3\text{H}]$ testosterone was measured—this assay has been validated as a specific measure of aromatization in granulosa cells⁵. Treatment of the cells with FSH shortly after plating in culture resulted in elevated aromatase activity after 24 h of treatment and a further increase after 48 and 72 h of treatment (Fig. 1). This effect of FSH confirmed earlier indications that it was involved in activating the aromatase system in immature granulosa cells, an important step leading to the formation of the preovulatory follicle⁶.

To investigate whether prolactin can influence this effect of FSH, prolactin was added alone, or together with FSH, to cultured cells. The amount of $^3\text{H}_2\text{O}$ released by the cell monolayers during a 2-h incubation with $[1\beta\text{-}^3\text{H}]$ testosterone was measured after 24, 48 and 72 h of treatment. Prolactin did not affect FSH-induced aromatase activity for the first 24 h of treatment, but between 24 and 48 h it inhibited the effect of FSH; this suppressed level of activity was also observed after 72 h (Fig. 1). In other experiments, concentrations of prolactin

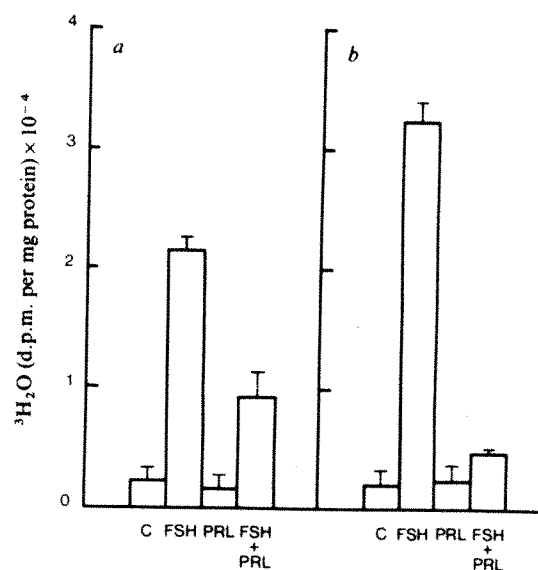


Fig. 2 Effects of FSH and prolactin on the aromatase activity of sonicates of ovarian granulosa cells using an optimal cell-free assay. Granulosa cells were prepared from diethylstilboestrol-primed immature rats and were cultured in 35 × 15 mm culture dishes in 1 ml Eagle's minimum essential medium for 48 (a) or 72 (b) h in the presence or absence of 300 ng NIH-FSH-S13 ml^{-1} and/or 1 μg NIH-P-S-7 ovine prolactin ml^{-1} . Medium was changed and treatments repeated after each 24-h culture period. At 48 and 72 h, cultures from each treatment group were washed, cells were scraped from the plates and transferred to tubes in 20 mM phosphate buffer, pH 7.4. The cell suspensions were sonicated on ice and 0.2-ml aliquots of the sonicates dispensed into assay tubes containing the following reagents: 0.1 ml of 20 mM phosphate buffer; 0.1 ml of buffer containing MgCl_2 and dithiothreitol to give a final concentration of 20 mM and 5 mM respectively; 50 μl NADPH generator (final concentration of NADP, 0.5 mM; glucose-6-phosphate, 20 mM; and glucose-6-phosphate dehydrogenase, 2U ml^{-1}); and 0.05 ml testosterone solution containing 0.4 μCi $[1\beta\text{-}^3\text{H}]$ testosterone and unlabelled testosterone to give a final concentration of 0.25 μM . Incubations with gentle shaking were for 1 h at 37 °C in an atmosphere of 95% O_2 and 5% CO_2 . The reaction was stopped by cooling tubes in ice-water and adding Norit A-activated charcoal (25 mg in 0.1 ml H_2O) to adsorb the steroids. Aliquots of the supernatant were counted in 5 ml Aquasol. Data are the means $\pm$ s.e.m. ($n = 3$).

from 10 to 100 ng ml^{-1} also inhibited the FSH-induced response. These concentrations are within the physiological range found in human follicular fluid⁷. No effects of prolactin alone were detectable during 72 h of culture, but this was not surprising as undifferentiated granulosa cells from immature animals do not contain specific prolactin binding sites⁸. Wang *et al.*⁹ showed that FSH stimulated the appearance of specific prolactin receptor sites on granulosa cells *in vitro* which may explain the 24-h lag before the effects of prolactin on FSH-induced aromatase activity were apparent.

To eliminate the possibility that prolactin influences aromatase activity by restricting the permeability of substrate or the availability of co-factors to the enzyme system, cells in culture were treated with hormones for various periods, after which cell sonicates were prepared and aromatase activity measured in the cell-free system in optimal conditions. As before⁵, pretreatment of granulosa cells in culture with FSH increased aromatase activity in the cell-free preparations (Fig. 2). Prolactin effectively reduced the amount of active aromatase induced by FSH during the first 48 h of culture and this effect increased after 72 h of treatment. Incubation of cell sonicates in optimal conditions showed that FSH and prolactin influenced aromatase activity in cultured granulosa cells by mechanisms which were independent of effects on the availability of substrate and

co-factor. FSH acts on granulosa cells to increase the amount of active aromatase, and prolactin seems to interfere with the ability of FSH to bring about this effect.

In addition to inhibiting FSH-induced oestrogen biosynthesis in rat granulosa cells, prolactin ($50\text{--}1,000\text{ ng ml}^{-1}$) affects steroidogenesis by stimulating progesterone synthesis⁹; the mechanism of two such different responses is intriguing. However, in human granulosa cells, high levels of prolactin seem to inhibit progesterone production¹⁰, and divergent effects have been demonstrated in porcine granulosa cells depending on the degree of differentiation¹¹.

When prolactin levels are high due to physiological (for example, lactation) or pathological (for example, hyperprolactinaemia) conditions, the inhibition of oestrogen biosynthesis and the concomitant changes in progesterone biosynthesis may establish a microenvironment in which the follicles can no longer develop normally and ovulation is suppressed¹.

We thank Heather McKeracher for technical assistance. This research was supported by the MRC of Canada.

Received 15 September 1980; accepted 2 February 1981.

- McNeilly, A. S. *Nature* **284**, 212 (1980).
- Besser, G. M., Parke, L., Edwards, C. R. W., Forsyth, I. A. & McNeilly, A. S. *Br. med. J.* **3**, 669 (1972).
- Frantz, A. G., Kleinberg, D. L. & Noel, G. L. *Recent Prog. Horm. Res.* **28**, 527 (1972).
- Engel, L. in *Handbook of Physiology* Vol. 2, Part 1 (eds Greep, R. O. & Astwood, E. B.) 467 (Am. Physiol. Soc., Washington DC, 1973).
- Gore-Langton, R. E. & Dorrington, J. H. *Molec. Cell Endocr.* (in the press).
- Dorrington, J. H., Moon, Y. S. & Armstrong, D. T. *Endocrinology* **97**, 1328 (1975).
- McNatty, K. P., Hunter, W. M., McNeilly, A. S. & Sawers, R. S. J. *Endocr.* **64**, 555 (1975).
- Richards, J. S. & Williams, J. J. *Endocrinology* **99**, 1571 (1976).
- Wang, C., Hsueh, A. J. W. & Erickson, G. F. *J. biol. Chem.* **254**, 11330–11336 (1979).
- McNatty, K. P., Sawers, R. & McNeilly, A. S. *Nature* **250**, 653 (1974).
- Velhuis, J. D., Klase, P. & Hammond, J. M. *Endocrinology* **107**, 42–46 (1980).

α -Lactalbumin is not a marker of human hormone-dependent breast cancer

Len Hall, Roger K. Craig*, Michael S. Davies,
David N. L. Ralphs & Peter N. Campbell

Courtauld Institute of Biochemistry, and the Department of Surgical Studies, The Middlesex Hospital Medical School, London W1P 7PN, UK

It has been proposed that milk protein gene expression in human breast carcinomas may indicate a functional oestrogen receptor mechanism, and may therefore be diagnostic of tumours responsive to endocrine therapy¹. Unfortunately this has not been proved, largely because of inconsistencies in the immunoassay procedures used to identify milk proteins, in particular α -lactalbumin, in tissue extracts or serum^{1–12}. Alternative procedures include the identification of milk protein mRNA by cell-free protein synthesis¹³, and the identification of milk protein RNA transcripts by hybridization to a sequence-specific probe. Here we describe experiments using α -lactalbumin cDNA probes, purified using recombinant plasmids¹⁴, which demonstrate that although polyadenylated and non-polyadenylated α -lactalbumin transcripts are present in normal human mammary tissue during pregnancy and lactation, α -lactalbumin transcripts are not detectable in the human tumour tissues studied. These experiments do however show that a peptide, which shares antigenic determinants with human α -lactalbumin is present in some breast tumour tissues.

* To whom correspondence should be addressed.

Human and guinea pig α -lactalbumin cDNA hybridization probes were used in RNA excess hybridization studies to compare the relative amounts of human α -lactalbumin RNA transcripts in different RNA preparations. Guinea pig α -lactalbumin cDNA, which fortuitously has extensive sequence homology with the human α -lactalbumin cDNA¹⁴, was also used as only limited amounts of human milk protein mRNA were available for production of cDNA probes. As can be seen in Fig. 1, >60% of guinea pig α -lactalbumin cDNA was protected from S_1 nuclease digestion when hybridized with poly(A)-containing RNA from lactating human mammary tissue 15 days post partum¹³. Furthermore, the kinetics of hybridization were identical to those of the homologous hybridization using human α -lactalbumin cDNA. Analysis of poly(A)-containing RNA

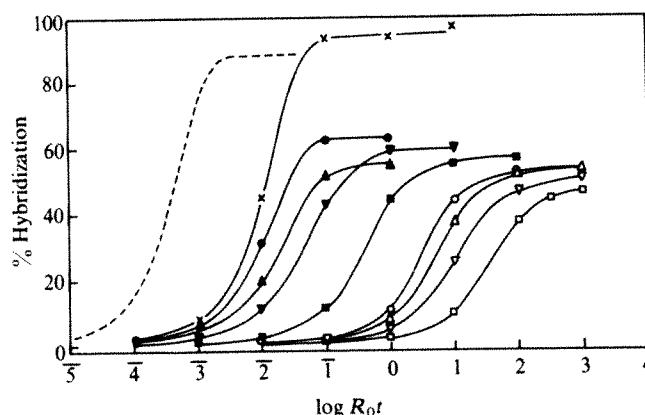


Fig. 1 Abundance of human α -lactalbumin transcripts in polyadenylated and non-polyadenylated RNA fractions isolated from normal mammary tissues of lactating and pregnant patients. Poly(A)-containing RNA was isolated from frozen human mammary tissue¹³ and hybridized either to human or guinea pig α -lactalbumin-specific cDNA as described elsewhere¹⁴. Non-polyadenylated RNA was defined as that which did not bind to oligo (dT)-cellulose at a salt concentration of 100 mM (ref. 15). The volume of individual hybridization assays (up to 100 μ l) and the RNA concentration (up to 5 mg ml⁻¹) were increased where appropriate to ensure an adequate RNA excess. Human α -lactalbumin cDNA versus poly(A)-containing RNA from 15-day post partum lactating human mammary tissue (x); guinea pig α -lactalbumin cDNA versus poly(A)-containing RNA from 15-d (●), 4-d (▲) and 330-d (▼) post partum mammary tissues, and normal breast tissue from a patient 32 weeks into pregnancy (■); guinea pig α -lactalbumin cDNA versus non-polyadenylated RNA from 15-d (○), 4-d (△) and 330-d (▽) post partum and 32-week pregnant (□) mammary tissue. The dashed line represents the homologous hybridization kinetics of highly purified rabbit globin mRNA standard¹⁴.

populations isolated from normal lactating mammary tissue 4 and 330 days post partum, and normal mammary tissue 32 weeks into pregnancy, using the guinea pig cDNA probe, demonstrated the presence of α -lactalbumin mRNA in each preparation but at different concentrations (Fig. 1). Compared with the 15-day post partum population, α -lactalbumin transcripts were present at 1.5-fold, 4-fold and 30-fold lower concentrations in the 4-day post partum, 330-day post partum and pregnant tissues, respectively. Analysis of the non-polyadenylated RNA from each tissue (RNA that did not bind to oligo (dT)-cellulose and contained polyadenylic acid tracts <35 nucleotides long¹⁵) showed that α -lactalbumin transcripts were present in all the preparations examined, again at different relative concentrations (Fig. 1). The overall analysis of normal human mammary tissue demonstrates that in common with the rat^{16,17}, mouse^{18,19} and guinea pig²⁰, α -lactalbumin gene

expression occurs in the mammary gland before parturition as well as throughout lactation. In common with many mRNA species including other milk protein mRNAs, human α -lactalbumin mRNA is distributed between the polyadenylated and non-polyadenylated RNA populations¹⁵.

Similar examination of polyadenylated and non-polyadenylated RNA populations isolated from human mammary tumours produced unexpected results. Poly(A)-containing RNA was isolated from 12 breast tumours (including adenocarcinomas, infiltrating duct carcinomas and a benign tumour): six were examined using a human cDNA probe, one of which, designated T2, an invasive duct carcinoma, had previously been shown to synthesize α -lactalbumin as determined by a combination of mRNA-directed cell-free protein synthesis, antibody precipitation and SDS-polyacrylamide gel electrophoresis¹³. However, hybridization studies (Fig. 2a) showed that none contained α -lactalbumin transcripts as there was only limited hybridization when compared with controls (chick oviduct and human placental poly(A)-containing RNA). The remaining six, which were analysed using the guinea pig cDNA probe, also contained no α -lactalbumin transcripts (Fig. 2b), even though three of these tumours had been independently shown by radioimmunoassay to contain significant amounts of α -lactalbumin (1,029, 716 and 1,554 pg per mg protein). Two of those containing α -lactalbumin as determined by radioimmunoassay also had elevated levels of oestrogen and progesterone receptors. Analysis of the non-polyadenylated RNA

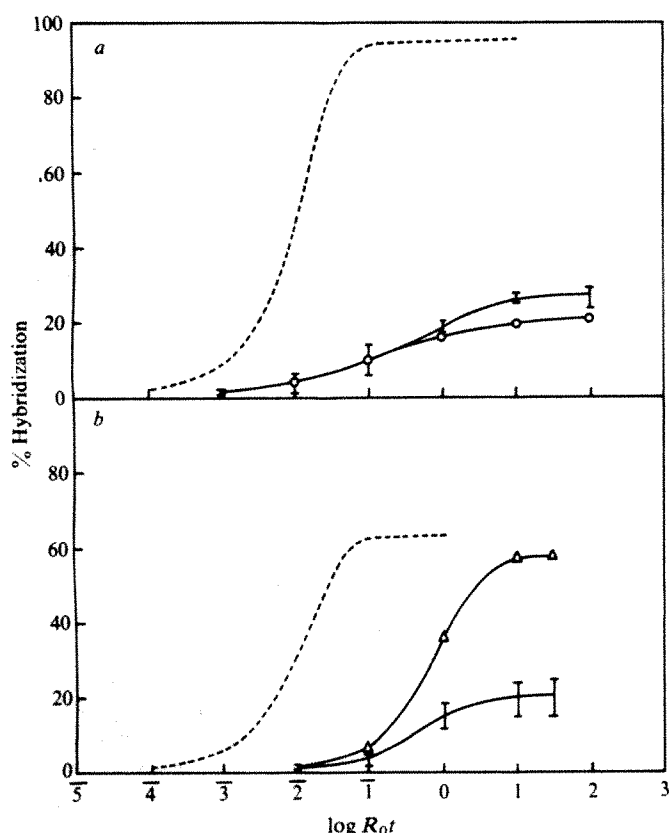


Fig. 2 Hybridization of poly(A)-containing RNA from human breast tumours to α -lactalbumin cDNA. Poly(A)-containing RNA was isolated from 12 frozen breast tumours¹³: six of these were challenged with (a), human α -lactalbumin cDNA and six with guinea pig α -lactalbumin cDNA (b). In each case the range of values obtained is indicated by error bars. Dotted lines represent the hybridization of poly(A)-containing RNA from 15-d postpartum lactating breast tissue to human (a) and guinea pig (b) α -lactalbumin cDNAs. Negative and positive controls were: (O), human α -lactalbumin cDNA versus human poly(A)-containing placental RNA; Δ , guinea pig α -lactalbumin cDNA versus a mixture of poly(A)-containing RNA from tumour and lactating tissue (99:1), respectively.

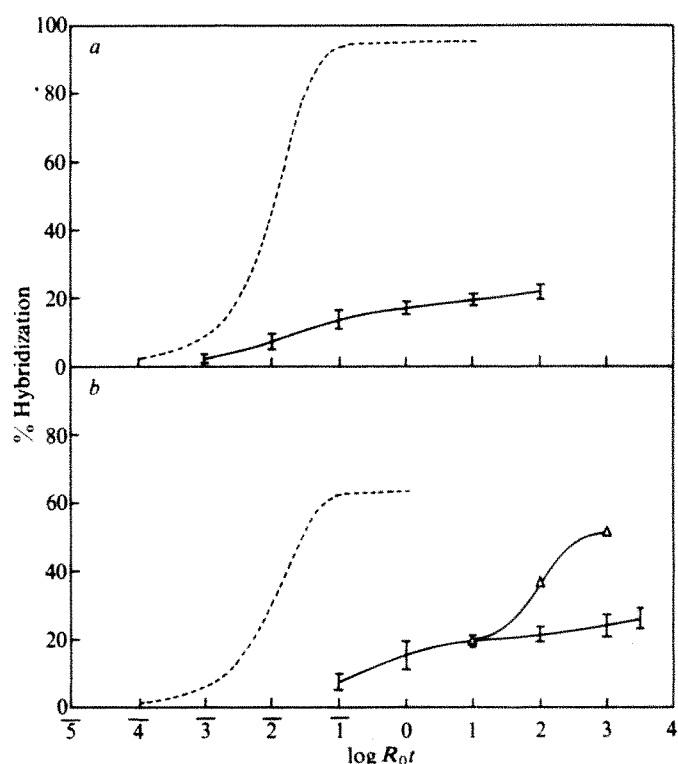


Fig. 3 Hybridization of poly(A)-containing and total cellular RNA from established tumour cell lines of human mammary origin, to α -lactalbumin cDNA. Twenty 9-cm-diameter dishes of semiconfluent cells were washed briefly with phosphate-buffered saline then lysed *in situ* with 10 mM Tris-HCl, pH 7.6, containing 1 mM EDTA, 2% (w/v) SDS, 0.5 mg ml⁻¹ proteinase K (1 ml per plate). The lysates were then processed as described for frozen tissue¹³. a, Human α -lactalbumin cDNA versus poly(A)-containing RNA from MCF-7 cells (both early and late passages, with and without added oestrogen²⁵), Hs578T cells and T-47D cells; error bars indicate range of values obtained. b, Guinea pig α -lactalbumin cDNA versus total cellular RNA from the above cell lines. Dotted lines represent the hybridization of poly(A)-containing RNA from 15-d postpartum lactating breast tissue to human (a) and guinea pig (b) α -lactalbumin cDNAs; Δ , guinea pig α -lactalbumin cDNA versus a mixture of MCF-7 total cellular RNA and polyadenylated RNA from 15-d postpartum lactating tissue (10,000:1).

from eight of the twelve tumours studied, including all those supposedly synthesizing α -lactalbumin, also proved to be negative (results not shown). As a positive control, poly(A)-containing RNA from 15-day lactating human tissue¹³ was mixed with a 100-fold excess of poly(A)-containing RNA from tumour tissue, then hybridized to guinea pig α -lactalbumin cDNA. This resulted in maximum hybridization (Fig. 2b) with the expected shift of hybridization kinetics (two logarithmic decades). It is unlikely that the absence of α -lactalbumin transcripts in tumour RNA preparations reflects degradation during extraction, as we estimated the mean size of the poly(A)-containing RNA to be 1,250 nucleotides in denaturing conditions, while cDNA prepared from the same RNA populations was of the same size, as determined by agarose gel electrophoresis in denaturing conditions.

We also analysed human cell lines of mammary origin: Hs578T²¹, T-47D²² and MCF-7²³ cell lines, all of which have been reported to synthesize either casein or α -lactalbumin²⁴. We found no evidence for the presence of α -lactalbumin transcripts, whether human α -lactalbumin cDNA was hybridized to poly(A)-containing RNA preparations (Fig. 3a) or whether guinea pig α -lactalbumin cDNA was hybridized to total cellular RNA (Fig. 3b). In the case of the MCF-7 cell line, RNA was analysed from early(15)- and late(275)-passage cells, and from cells stimulated with oestrogen²⁵, with negative results. As a

positive control, poly(A)-containing RNA from 15-day post partum lactating human mammary tissue was mixed with a 10,000-fold excess of MCF-7 total cellular RNA. This produced the expected degree of hybridization and shift in kinetics (Fig. 3b), thus confirming the specificity of the assay.

Overall it is apparent that α -lactalbumin gene expression does not occur to a measurable extent in any of the tumours or cell lines examined, using hybridization procedures which can detect as little as one polyadenylated, or <10 non-polyadenylated RNA transcripts per cell. The absence of α -lactalbumin RNA transcripts in the cell lines was not surprising, because although preliminary *R₀t* analysis and Southern blotting of cell-line and human placental restricted DNA demonstrated that the genes were present, we had been unable to demonstrate the presence of α -lactalbumin mRNA by cell-free protein synthesis (L.H. and R.K.C., unpublished data). However, the absence of α -lactalbumin transcripts from tumour tissue was unexpected, particularly as we have previously reported that a peptide with the expected electrophoretic mobility of pre- α -lactalbumin¹³ was synthesized in a cell-free protein synthesizing system directed by a tumour mRNA preparation, and that this peptide could be preferentially precipitated by α -lactalbumin antisera. This particular observation has proved reproducible using the same tumour (T2) mRNA and our own or an alternative source of α -lactalbumin antiserum. Because no α -lactalbumin mRNA was present in this tumour mRNA preparation, we conclude that both α -lactalbumin antisera preparations were reacting with a peptide which not only shared antigenic determinants with α -lactalbumin, but was also of the same apparent molecular weight as pre- α -lactalbumin.

Although we have not established the precise nature of the peptide(s) precipitated by antiserum to human α -lactalbumin, we consider that the close correlation between ' α -lactalbumin' and oestrogen receptors, reported to indicate a functional oestrogen receptor mechanism¹, may still prove useful once this peptide has been identified. Consequently we are screening our library of human mammary cDNA plasmids for those plasmids containing sequences that are abundant in tumour mRNA populations, several recombinants are being studied. Thus although the results are negative from a clinical point of view, we have demonstrated that sequence-specific cDNA hybridization probes generated using recombinant DNA techniques may soon be important as diagnostic tools in medicine. To conclude, using highly sensitive and specific hybridization techniques, we found no evidence to support the concept that α -lactalbumin will prove to be a general marker for human hormone-dependent breast cancer.

We thank the Cancer Research Campaign for supporting this work and Dr A. Howell and his colleagues for performing the α -lactalbumin radioimmunoassay, the hormone receptor analyses, and providing an alternative source of human α -lactalbumin antisera. Avian myeloblastosis virus reverse transcriptase was provided by Drs J. Beard and J. Gruber. Human mammary cell lines Hs578T and T-47D, and early- and late-passage MCF-7 cell lines were given by Drs Adeline Hackett, I. Keydar and C. McGrath, respectively.

Received 12 January; accepted 26 January 1981.

1. Woods, K. L., Cove, D. H. & Howell, A. *Lancet* **ii**, 14-16 (1977).
2. Woods, K. L., Cove, D. H., Morrison, J. M. & Heath, D. A. *Eur. J. Cancer* **15**, 47-51 (1979).
3. Kleinberg, D. L., Todd, J. & Groves, M. L. *J. clin. Endocr. Metab.* **45**, 1238-1250 (1977).
4. Kleinberg, D. L. & Todd, J. *Cancer Res.* **38**, 4318-4322 (1978).
5. Kleinberg, D. L. *Science* **190**, 276-278 (1975).
6. Rose, H. N. & McGrath, C. M. *Science* **190**, 673-675 (1975).
7. Schultz, G. S. & Ebner, K. E. *Cancer Res.* **37**, 4489-4492 (1977).
8. Monaco, M. E., Bronzert, D. A., Tormey, D. C., Waalkes, P. & Lippman, M. E. *Cancer Res.* **37**, 749-754 (1977).
9. Hendrick, J. C. & Franchimont, P. *Eur. J. Cancer* **10**, 725-730 (1974).
10. Woods, R. L., Searle, F., Wilson, P., Bagshawe, K. D. & Cooper, E. H. *Eur. J. Cancer*, **13**, 1113-1117 (1977).
11. Pick, A., Bussolati, G. & Di Carlo, F. *J. natn. Cancer Inst.* **58**, 1483-1484 (1977).
12. Stevens, V., Lawrence, D. J. R. & Ormerod, M. G. *Clin. chim. Acta* **87**, 144-157 (1978).
13. Hall, L., Craig, R. K. & Campbell, P. N. *Nature* **277**, 54-56 (1979).
14. Hall, L., Davies, M. S. & Craig, R. K. *Nucleic Acids Res.* **9**, 65-84 (1981).
15. Bathurst, I. C., Craig, R. K., Herries, D. G. & Campbell, P. N. *Biochem. J.* **192**, 489-498 (1980).

16. Nardacci, N. J., Lee, J. W. C. & McGuire, W. L. *Cancer Res.* **38**, 2694-2699 (1978).
17. Nakhachi, H. L. & Qasba, P. K. *J. biol. Chem.* **254**, 6016-6025 (1979).
18. Ono, M. & Oka, T. *Cell* **19**, 473-480 (1980).
19. Ono, M. & Oka, T. *Science* **207**, 1367-1369 (1980).
20. Burditt, L. J., Parker, D., Craig, R. K., Getova, T. & Campbell, P. N. *Biochem. J.* **194**, 999-1006 (1981).
21. Hackett, A. J. *et al. J. natn. Cancer Inst.* **58**, 1795-1806 (1977).
22. Keydar, I. *et al. Eur. J. Cancer* **15**, 659-670 (1979).
23. Soule, H. D., Vazquez, A., Long, A., Albert, S. & Brennan, M. A. *J. natn. Cancer Inst.* **51**, 1409-1413 (1973).
24. Engel, L. W. & Young, N. A. *Cancer Res.* **38**, 4327-4339 (1978).
25. Lippman, M. E. & Bolan, G. *Nature* **256**, 592-593 (1975).

Enhanced frequency of chromosome aberrations in lymphocytes of male compared with female muntjacs after X-ray irradiation *in vitro*

B. C. Das & T. Sharma

Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi-221005, India

Recently, there has been considerable interest in the study of radiation-induced dicentrics and other chromosome aberrations in phytohaemagglutinin (PHA)-stimulated blood lymphocytes of various mammals, including man, with the aim of establishing a proper biological dosimetry for the assessment of genetic radiation hazards in human. These studies have revealed that the radiosensitivity of lymphocytes differs between species and even between individuals of the same species¹⁻⁵, but the cause(s) of these variations still remains unclear. The life-shortening response to whole-body X-ray irradiation is influenced by both age and sex of the individuals⁶⁻⁹, but although age has been shown to have effects on chromosome aberration yields^{4,10-12} *in vitro*, the influence of sex on the sensitivity of lymphocyte chromosomes to radiation-induced damage has not been reported. It is generally thought that sex does not influence the aberration yield within the limitation of measurements used in such studies. However, we report here a study of X-ray induced chromosome aberration in lymphocytes of Indian muntjac, in which the yield of dicentrics was found to be consistently greater in male than in female lymphocytes.

We prepared lymphocyte cultures from healthy, young adult muntjacs (*Muntiacus muntjak*) selected from a small breeding colony which we maintained. The specimens were chosen so that we could do experiments on blood samples of three pairs of muntjacs of similar age and body weight. This barking deer has only six chromosomes in the female, with XX sex chromosomes, and seven in the male, with XY₁Y₂ sex chromosomes^{13,14}, and both the complements are distinguishable from each other by distinct constitution and morphology of the sex chromosomes (Fig. 1 a-d). This unique sex-based karyotypic dimorphism was used in a 'mixed blood culture' (MBC) method where whole blood drawn from males and females was mixed in equal proportion and then irradiated with various doses of X rays (110 kV, 4 mA, 1 mm Al and 100 rad min⁻¹ at 37°C) and cultured. This method ensures greater accuracy as it eliminates the possibility of errors inherent in irradiating and growing cultures separately for male and female individuals. The frequent metaphases for both sexes in MBCs, and the relative frequencies of first-, second- and third-division cells obtained at different times of fixation from male and female blood cultures set up separately, or after mixing, indicated that the normal progression of cell division had no obvious 'cross-reaction' in the MBCs. The results will be published in detail elsewhere.

A low concentration (0.1 µg ml⁻¹) of 5-bromodeoxyuridine (BUdR)¹⁵ was added at the start of culture and the chromosome preparations were stained using the Hoechst-sunlight-Giemsa technique^{16,17} to score aberrations only from the first-cycle

Table 1 Frequencies of dicentrics and other aberrations in first-division metaphases of male and female muntjac lymphocytes in MBCs after X-ray irradiation

Muntjac pairs	Dose (rad)	Sex	Fixation time (h)	Cells scored	Normal cells	Dicentrics $\pm$ fragments	Rings $\pm$ fragments	Breaks	Gaps	Deletions (TD + ID)
1 Male (Tomy) Female (Sindhu)	0	M	32	300	289	0	0	1	6	9
		F		300	293	0	0	0	3	10
	100	M	33	400	250	27*	2	11	40	89*
		F		400	278	14	7	9	43	61
	200	M	34	400	153	104†	11	17	43	210†
		F		400	199*	46	13	13	31	150
	300	M	35	300	62	128†	16	19	35	244
		F		300	87*	63	24	12	30	206
	400	M	36	200	15	117†	18	11	19	271
		F		200	29*	77	16	13	26	259
	0	M	32	200	192	0	0	1	2	5
		F		200	194	0	0	1	1	4
2 Male (Homy) Female (Bindu)	100	M	33	300	216	25*	6	8	11	34
		F		300	185	12	7	4	13	25
	200	M	34	200	101	39*	4	3	7	68
		F		200	117	22	4	4	9	55
	300	M	35	100	20	39*	5	12*	8	56
		F		100	24	27	6	4	9	79*
	0	M	32	200	195	0	0	0	2	3
		F		200	196	0	0	0	1	3
	100	M	33	300	187	25*	3	18*	20*	77†
		F		300	282†	13	4	7	3	42
	200	M	34	200	85	41†	11	20	18†	64
		F		200	120†	19	4	14	2	64
3 Male (Sony) Female (Rindu)	300	M	35	225	52	84†	19	12	19	177*
		F		225	85†	50	18	14	14	142
	400	M	36	100	10	56*	12	13	6	172†
		F		100	13	35	7	9	7	117

TD, terminal deletion; ID, interstitial deletion.

* Significant in χ^2 test, $P < 0.05$.† Significant in χ^2 test, $P < 0.01$.**Table 2** Frequencies of dicentrics and other aberrations in first-division metaphases of male and female muntjac lymphocytes in MBCs fixed at 40 h after X-ray irradiation

	Dose (rad)	Sex	Fixation time (h)	Cells scored	Normal cells	Dicentrics $\pm$ fragments	Rings $\pm$ fragments	Breaks	Gaps	Deletions (TD + ID)
Male (Tomy)	0	M	40	100	87	0	0	0	6	7
		F		100	90	0	0	0	9	2
Female (Sindhu)	100	M	41	280	169	27*	2	17	29	48
		F		280	179	12	8	10	38	55
200		M	42	310	127	62†	14*	15	34	126
		F		310	144	33	4	20	24	127
300		M	43	200	40	90†	15	5	15	148
		F		200	53	48	11	7	17	122
400		M	44	150	11	84*	19	8	16	205
		F		150	21	57	17	8	17	172

TD, terminal deletion; ID, interstitial deletion.

* Significant in χ^2 test, $P < 0.05$.† Significant in χ^2 test, $P < 0.01$.**Table 3** Frequencies of dicentrics and other aberrations in first-division metaphases from cultures set up separately for male and female muntjac blood after X-ray irradiation

Munjac pair	Dose (rad)	Sex	Fixation time (h)	Cells scored	Normal cells	Dicentrics $\pm$ fragments	Rings $\pm$ fragments	Breaks	Gaps	Deletions (TD + ID)
Male (Tomy)	0	M	32	150	146	0	0	0	2	3
		F		150	147	0	0	0	2	2
Female (Sindhu)	100	M	33	150	99	13*	3	14*	13*	24
		F		150	117	4	1	4	3	22
200		M	34	250	93	41*	15	15	12	112
		F		250	131*	24	8	18	18	94
300		M	35	150	32	57*	11	22	11	119
		F		150	43	37	16	12	8	138
400		M	36	150	12	73*	23	13*	14	172*
		F		150	15	49	27	4	12	135

TD, terminal deletion; ID, interstitial deletion.

* Significant in χ^2 test, $P < 0.05$.



Fig. 1 Metaphase chromosomes of male and female muntjacs. *a*, First-cycle and *b*, second-cycle metaphase preparations from female muntjac lymphocytes showing six chromosomes with XX. *c*, First-cycle and *d*, second-cycle metaphase preparations from male muntjac lymphocytes showing seven chromosomes with XY_1Y_2 , a multiple sex chromosome system due to X-autosome translocation.

metaphases^{3,18-21}. For multiple sampling, fixation was done at 32 and 40 h to analyse aberrations. The frequencies of dicentric and other chromosome aberrations were scored from coded slides from 50–100 first-cycle male and female metaphases. The frequency of sister chromatid exchanges (SCEs) was scored from second-cycle metaphases and the experiments were repeated two to three times for each pair of muntjac.

The results obtained from the MBCs fixed at 32 h (Table 1) and 40 h (Table 2) as well as from the cultures of male and female lymphocytes set up separately (Table 3) indicated that the male lymphocytes were consistently more radiosensitive than female lymphocytes, as they showed a significantly higher (1.86 ± 0.25 , $P < 0.01$ or 0.05) frequency of dicentric (Fig. 2). Although the frequency of deletions, which did not include fragments associated with dicentric and rings, was quite high and at times showed significant variation between the two sexes, the frequency was not significantly different in most cases. With regard to symmetrical exchanges, there was sometimes a slight increase in the yield in females but the frequency was found to be statistically insignificant when compared with yields in males. The analysis was not done after banding of chromosomes, therefore their frequencies are not very precise the data on symmetrical translocations have not been shown. The yield of dicentric in cultures set up separately for male and female was slightly lower than in the MBCs (Fig. 2) but the magnitude of difference in the yield of dicentric between the two sexes still

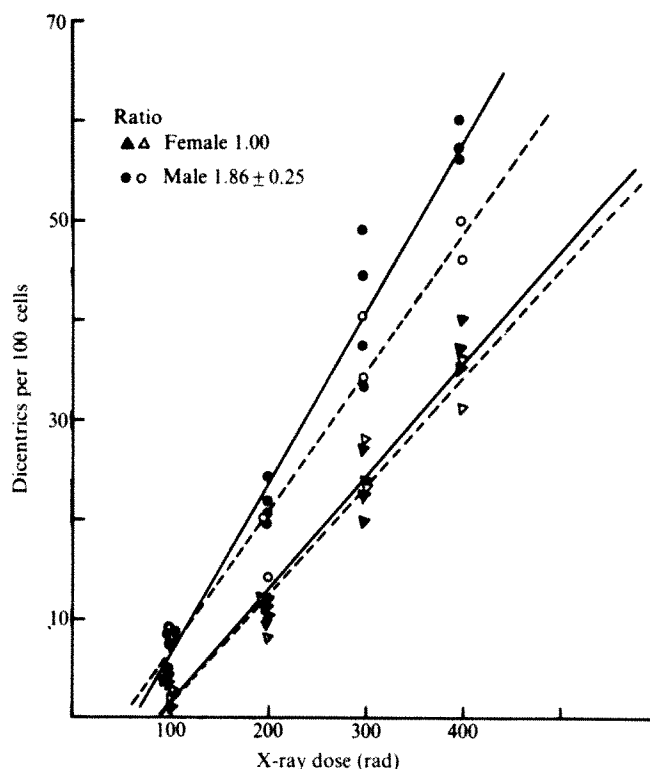


Fig. 2 Dose-response curves for dicentric in male and female muntjac lymphocytes grown in MBCs (●, male; ▲, female) and separately (○, ♂ male; △, ♀ female), after exposure to various doses of X rays. The curves are the least-squares regression lines of $y = a + bD$ ($r = 0.98$, $P < 0.001$).

remained the same. The frequency of other chromosome aberrations (Tables 1, 2, 3) however, did not show any significant difference between the two sexes, either in the MBCs or in the separate cultures. The frequency of SCEs was the same in male and female lymphocytes and was of the order of those occurring spontaneously, even after exposure to 400 rad (Table 4). This corroborates the suggestion that the mechanisms of production of SCEs and chromosome aberrations are different²⁰⁻²³.

Our previous study¹⁸ using the BUdR-harlequin technique on male muntjac lymphocytes fixed at 40 and 48 h after X-ray irradiation and initiation of cultures, and the present study of MBCs fixed at 32 and 40 h (Tables 1, 2) gave similar yields of dicentric and rings in the first-cycle metaphases. This strongly suggests homogeneous sensitivity of muntjac lymphocytes to X-ray irradiation, which has also been reported for humans^{3,21}. Nevertheless, there are reports on human lymphocytes which suggest the presence of at least two subpopulations that differ in rate of progression of cell cycle and radiosensitivity^{20,24,25}. In view of the conflicting reports on human lymphocyte radiosensitivity, we scored dicentric in lymphocytes of both sexes of muntjac only from the first wave of mitosis, mainly at 32 h, even though the yield of metaphases was very low at this hour, particularly in the case of male metaphases.

The significant difference observed in the frequency of X-ray induced dicentric therefore indicates differential radiosensitivity of lymphocytes in the two sexes of muntjac. The low number of chromosomes in muntjac may suffer from 'distortion' in radiation-induced chromosome exchange analysis²⁶ but the slight difference in the chromosome number of males and females of muntjac is not expected to influence the yield of dicentric in the two sexes. Thus these results reflect a real difference in radiosensitivity. Differential repair efficiency and/or other intrinsic factors may be responsible for this differential sensitivity. It is, however, interesting to note that in

Table 4 Sister chromatid exchange (SCE) frequency in control and irradiated cultures of male and female muntjac lymphocytes

X-ray dose (rad)	Sex	Total metaphases scored	Total SCEs scored	SCEs per cell $\pm$ s.e.
0	Male	50	363	7.26 ± 0.46
	Female	50	365	7.30 ± 0.42
100	Male	50	382	7.64 ± 0.72
	Female	50	448	8.96 ± 1.37
200	Male	50	341	6.82 ± 0.29
	Female	50	325	6.50 ± 0.59
400	Male	50	362	7.24 ± 0.39
	Female	50	382	7.64 ± 0.82

Down's syndrome, the presence of an extra chromosome 21 has been shown to cause a 50–100% increase in the incidence of radiation-induced exchange-type aberrations (dicentric and rings), whereas the frequencies of breaks and other intra-chromosomal changes are not affected^{10,11,27}. Although the situation is not parallel, we could assume that the presence of Y in male muntjac as an extra chromosome in the complement in some way accounts for the higher yield of dicentric. Experiments on other species with similar sex chromosome constitution, resulting in a difference in chromosome number between the two sexes, may elucidate this point.

We thank Professor J. R. K. Savage for help with the manuscript, and Drs H. Loewe and V. V. Menon for the gift of Hoechst 33258 and statistical advice, respectively. The work was supported by grants from the International Atomic Energy Agency, contract nos 2126/RB and 2126/RI/RB and the University Grants Commission, India.

Received 16 December 1980; accepted 26 January 1981.

- Shankaranarayanan, K. *Mutat. Res.* **35**, 371–386 (1976).
- Preston, R. J. & Brewen, J. G. in *Mutagen Induced Chromosome Damage in Man* (eds Evans, H. J. & Lloyd, D. L.) 33–40 (Edinburgh University Press, 1978).
- Scott, D. & Lyons, C. Y. *Nature* **278**, 756–758 (1978).
- Sasaki, M. S., Tonomura, A. & Matsubara, S. *Mutat. Res.* **10**, 617–633 (1970).
- Bockhov, N. P. & Pilosov, R. A. *Genetics* **4**, 144–152 (1968).
- Holland, J. M. & Mitchell, T. J. *Radiat. Res.* **66**, 363–372 (1976).
- Abrams, H. L. *Proc. Soc. exp. Biol. Med.* **76**, 729–732 (1951).
- Lindop, P. J. & Rotblat, J. *Br. J. Radiol.* **35**, 23–31 (1962).
- Leonard, A., Maisin, J. R. & Deknigt, G. *Strahlentherapie* **137**, 92–100 (1969).
- Sasaki, M. S. & Tonomura, A. *Jap. J. hum. Genet.* **14**, 81–92 (1969).
- Evans, H. J. & Adams, A. in *Advances in Radiation Research, Proc. 4th int. Congr. Radiat. Res.* (eds Duplan, J. F. & Chapiro, A.) (Gordon & Breach, New York, 1973).
- Konopliannikova, O. A. *Genetika* **10**, 111–119 (1966).
- Wurster, D. H. & Benirschke, K. *Science* **168**, 1364–1366 (1970).
- Sharma, T. *Cytogenetics* **11**, 1–6 (1972).
- Sharma, T. & Das, B. C. *Mutat. Res.* (in the press).
- Perry, P. & Wolff, S. *Nature* **251**, 156–158 (1974).
- Goto, K., Akematsu, T., Shimazu, H. & Sugiyama, T. *Chromosoma* **53**, 223–230 (1975).
- Das, B. C. & Sharma, T. *3rd a. scient. Meet. of Environmental Mutagen Society of India*, Ahmedabad, Abst. 8 (1977).
- Savage, J. R. K. *Nature* **278**, 687 (1979).
- Bianchi, N. O., Bianchi, M. S. & Larramendy, M. *Mutat. Res.* **63**, 317–324 (1979).
- Leonard, A. & Decat, G. *Can. J. Genet. Cytol.* **21**, 473–478 (1979).
- Wolff, S., Rodin, B. & Cleaver, J. E. *Nature* **265**, 347–348 (1977).
- Wolff, S. A. *Rev. Genet.* **11**, 183–201 (1977).
- Bender, M. A. & Brewen, J. G. *Mutat. Res.* **8**, 383–399 (1969).
- Steffen, J. & Michalowski, A. *Mutat. Res.* **17**, 367–376 (1973).
- Savage, J. R. K. & Papworth, D. G. *J. theor. Biol.* **22**, 493–514 (1969).
- Kucerova, M. *Acta radiol.* **6**, 441–448 (1967).

Altered ribosomal RNA genes in mitochondria from mammalian cells with chloramphenicol resistance

Stephen E. Kearsey & Ian W. Craig

Genetics Laboratory, Department of Biochemistry,
University of Oxford, Oxford OX1 3QU, UK

Chloramphenicol resistance in mammalian cells is cytoplasmically inherited^{1,2}. In yeast, a similar phenotype is caused by mutations in the mitochondrial DNA (mtDNA), and sequencing of carefully constructed strains has identified nucleotide monosubstitutions in the 3' region of the large (21S) rRNA gene which correlate with the antibiotic resistance³. We have sequenced the corresponding section of mammalian mtDNA from chloramphenicol-resistant cell lines for comparison with the wild-type sequence. Differences between the sequences occur at positions similar to those altered in the yeast mutants, in a highly conserved region of the large (16S) rRNA gene.

The chloramphenicol-resistant cell lines investigated were 3T3CAP^R (ref. 4 and R. L. Siegel, unpublished data) derived from the mouse cell line 3T3, and the human line MC63 derived from HeLa B (ref. 5). Cytoplasm-cell fusion experiments have shown convincingly that the chloramphenicol resistance of these

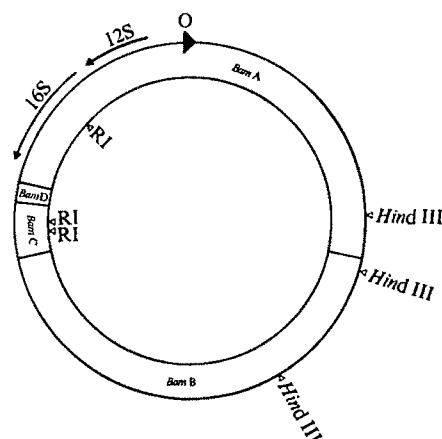


Fig. 1 The positions of *Bam*HI, *Eco*RI and *Hind*III restriction enzyme sites in mouse mtDNA²⁰; O denotes the H-strand replication origin, and the direction of transcription of the rRNA genes is marked²¹. The *Bam*HI-A fragment was purified by cloning in pBR322 (ref. 7), and the smaller *Bam*HI-*Eco*RI subfragment was isolated and digested with *Sau*3A. This was ligated into the double-stranded form (RFI) of a bacteriophage M13mp2 derivative adapted to accept *Bam*HI fragments^{9–11} (amber mutation in phage; G. Winter, unpublished results). Recombinant phage, which give colourless plaques on a lawn of *E. coli* JM101 (ref. 22), were grown up for screening as described in the text. The universal primer from pSP14 (ref. 23) was used for dideoxy sequencing¹⁴.

cell lines may be cytoplasmically transmitted to sensitive cells^{2,6}, although further genetic characterization has not been completed.

MtDNA from 3T3CAP^R cells was cloned as *Bam*HI restriction fragments using the plasmid pBR322 (Fig. 1 and ref. 7). Similar banks were also made using mtDNA from the parental cell line 3T3 and mouse liver. Plasmids containing the *Bam*HI-A fragment were prepared, allowing the purification of the smaller (1.3-kilobase) *Eco*RI-*Bam*HI subfragment from the plasmid insert. This contains the 3' end of the large (16S) rRNA gene (Fig. 1). The fragment was cut multiply with *Sau*3A for subcloning, using a bacteriophage M13mp2 derivative as vector^{8–10}.

Recombinant M13 clones from the mouse liver mtDNA bank were screened by partial sequencing reactions, using one dideoxynucleotide, ddT¹¹. Clones chosen for detailed sequence analysis were those containing restriction fragments having homology with the yeast 21S rRNA gene region in which nucleotide changes conferring chloramphenicol resistance are found³. As human and mouse mtDNA show high sequence homology in the rRNA genes¹², this screening was made easier by aligning the human 16S rRNA gene sequence¹³ with the yeast sequence. Comparison of the thymidine partial sequences of the mouse *Sau*3A fragments with the human sequence allowed the identification of several hybrid phage containing a 92-base pair fragment which spans the relevant region of the 16S rRNA gene. Two of these clones, containing this fragment in opposite orientations, were used to select clones from the 3T3 and 3T3CAP^R banks for dideoxy sequencing¹⁴. In this hybridization screening¹⁵, single-stranded DNA from clones containing the same restriction fragment as the reference clone, but in an opposite orientation, formed a hybrid with the reference clone DNA. These DNA hybrids have a lower mobility on agarose gels than the non-annealed single strands.

The MC63 mtDNA was digested with *Sau*3A and cloned directly into bacteriophage M13. Hybridization screening of 40 recombinant phages identified three homologous clones which were used for sequence determination.

Comparison of the mtDNA sequences obtained shows that the 92-base pair *Sau*3A fragments from 3T3CAP^R and 3T3 mtDNA differ by one nucleotide at position 30 (Fig. 2).

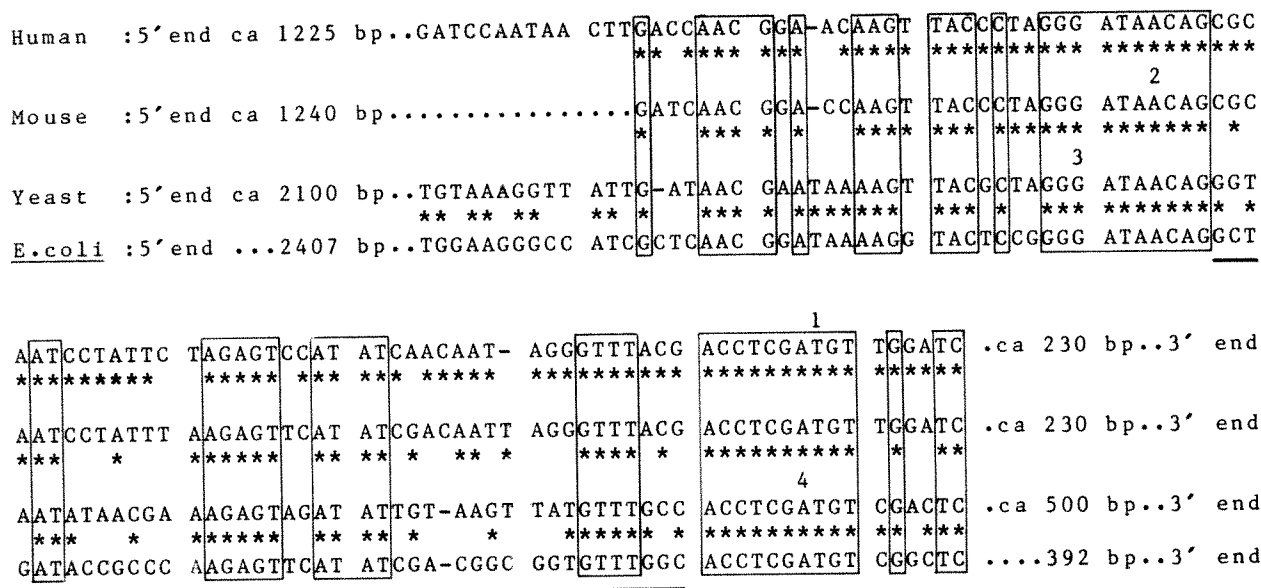


Fig. 2 Comparison of the DNA sequences in the 3' region of the large rRNA genes, showing positions of nucleotide changes in mtDNA from chloramphenicol-resistant cells. Sequences of the noncoding strand are given. The sequences compared are from human (placental) mtDNA¹³, mouse (3T3 or liver) mtDNA, yeast (*Saccharomyces cerevisiae* strain 1L9-8A/D122, with the C₃₂₃ site in its wild-type configuration) mtDNA³ and *E. coli* (*rrnB* gene)¹⁶. The sequences are aligned to show major blocks of homology, with gaps where appropriate. Asterisks show identity between adjacent sequences and boxed-in regions indicate sequences conserved between all four. Underlining indicates the position of the stem of a hairpin loop structure which is conserved between all four sequences. Nucleotide monosubstitutions in mtDNA sequences from chloramphenicol-resistant cells are indicated as follows: 1, MC63 (T to C); 2, 3T3CAP^R (A to T); 3, C₃₂₃ mutation in yeast (G to A)³; and 4, C₃₂₁ mutation in yeast (A to C)³.

Comparison with the yeast sequence data shows that this A-T transversion is only four nucleotides away from a homologous site at which a G-A transition in the *rib-1-B* locus causes the chloramphenicol resistance mutation C₃₂₃^R (ref. 3). This suggests that the sequence alteration is responsible for the antibiotic resistance phenotype of 3T3CAP^R although no functional genetic test is available. It is less likely that this mutation represents a silent polymorphism as the wild-type sequence compared is from the parental cell line. In any case, there are no differences between the 3T3 and mouse liver sequences (see Fig. 2) which suggests that polymorphism in the 3' region of the 16S rRNA gene is not high, and that the cloning steps used do not mutate the original nucleotide sequences.

The 104-base pair *Sau3A* fragment from MC63 mtDNA also shows one difference compared with the human placental sequence reported by Eperon *et al.*¹³ (Fig. 2). This monosubstitution, a T-C transition at position 96, is one nucleotide away from a homologous site in yeast mtDNA where an A-C transversion at the *rib-1-A* locus causes the chloramphenicol resistance mutation C₃₂₁^R (ref. 3).

The single nucleotide changes in the mtDNA from both the mouse and human mutant cell lines occur in short sequence blocks which are identical in the yeast mitochondrial 18S rRNA gene and the *Escherichia coli* 23S rRNA (*rrnB*) gene (Fig. 2 and refs 3, 16). Such conservation probably reflects functional constraints imposed on this region of the genome and suggests that a small genetic change, such as a base pair alteration, could have a considerable effect on the cell phenotype. Furthermore, this evolutionary stability makes it less likely that irrelevant polymorphisms are being detected.

Structural studies of the *E. coli* ribosome indicate that in the rRNA, the sequence shown in Fig. 2 occurs at the interface between the 30S and 50S subunits, the RNA being protected from kethoxal modification on subunit association¹⁷. This agrees well with the high degree of sequence conservation of this region through evolution compared with other parts of the large rRNA¹³. The nucleotides which are altered in chloramphenicol-resistant mutants are probably implicated in binding the 3' terminus of aminoacyl tRNA, because chloramphenicol blocks

this site¹⁸. A hairpin loop can be constructed between the mutation sites, and this feature is conserved despite local divergence of the primary sequences compared in Fig. 2 (ref. 19). If this structure occurs in the ribosome it would bring the putative chloramphenicol-binding sites together.

Further sequence analysis of mtDNA from other chloramphenicol-resistant cell lines should help to confirm the correlation suggested here. Chloramphenicol resistance is the only readily obtainable genetic marker in the mammalian cytoplasm, and a simple molecular basis for the mutation might help to explain why this is the case.

We thank Greg Winter for the primer, and Elin Munro for culturing and phenotypic characterization of cell lines. This work was supported by grants from the CRC and the Christopher Welch Trustees.

Received 2 December 1980; accepted 11 February 1981.

- Bunn, C. L., Wallace, D. C. & Eisenstadt, J. M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1681-1685 (1974).
- Munro, E., Siegel, R. L., Craig, I. W. & Sly, W. S. *Proc. R. Soc. B* **201**, 73-85 (1978).
- Dujon, B. *Cell* **20**, 185-197 (1980).
- Siegel, R. L. thesis, Univ. Washington (1977).
- Siegel, R. L., Jeffreys, A. J., Sly, W. & Craig, I. W. *Expl. Cell Res.* **102**, 298-310 (1976).
- McBurney, M. W. & Strutt, B. *Expl. Cell Res.* **124**, 171-180 (1979).
- Kearsey, S. E., Flanagan, J. G. & Craig, I. W. *Gene* **12**, 249-255 (1980).
- Messing, J., Gronenborn, B., Müller-Hill, B. & Hofschneider, P. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3642-3646 (1977).
- Gronenborn, B. & Messing, J. *Nature* **272**, 375-377 (1978).
- Rothstein, R. J., Lau, L. F., Bahl, C. P., Narang, S. A. & Wu, R. *Meth. Enzym.* **68**, 98-109 (1980).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
- de Francesco, L. & Attardi, G. *J. molec. Biol.* **139**, 85-93 (1980).
- Eperon, I. C., Anderson, S. & Nierlich, D. P. *Nature* **286**, 460-467 (1980).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Winter, G. & Fields, S. *Nucleic Acids Res.* **8**, 1965-1974 (1980).
- Brosius, J., Dull, T. J. & Noller, H. F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 201-204 (1980).
- Herr, W. & Noller, H. F. *J. molec. Biol.* **130**, 421-432 (1979).
- Nierhaus, D. & Nierhaus, K. H. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2224-2228 (1973).
- Baer, R. J. & Dubin, D. T. *Nucleic Acids Res.* **9**, 323-337 (1981).
- Moore, K. H., Johnson, P. H., Chandler, S. E. W. & Grossman, L. I. *Nucleic Acids Res.* **4**, 1273-1288 (1977).
- Batley, J. & Clayton, D. A. *Cell* **14**, 143-156 (1978).
- Messing, J. *Recombinant DNA Tech. Bull.* **2**, 43-48 (1979).
- Anderson, S., Gait, M. J., Mayol, L. & Young, I. G. *Nucleic Acids Res.* **8**, 1731-1743 (1980).

MATTERS ARISING

Sexual dimorphism in early anthropoids

FLEAGLE *ET AL.*¹ provide an estimate of sexual dimorphism in extinct anthropoid primates based on a positive relationship between the degree of sexual dimorphism in body weight and the coefficient of variation (c.v.) of mandible depth in extant anthropoids. A cursory glance at Table 1 shows that their data on the degree of body weight dimorphism are incorrect. As a striking example, capuchin monkeys (*Cebus apella*) are supposedly more dimorphic than gorillas (*Gorilla gorilla*) which, as any visitor to the zoo knows, is simply not the case. Gorillas are far more dimorphic than capuchin monkeys.

A closer examination of the data reveals that 9 (those taken from our study on sexual dimorphism²) out of 13 species do not express a mean male weight/mean female weight ratio as intended by Fleagle *et al.*, but instead show a ratio between the cube roots of male and female weights. In our study we used the cube root of body weight to obtain a linear dimension comparable to a linear canine dimension. Using corrected data for body weight, I recalculated its correlation with the c.v. of mandible depth and obtained a correlation coefficient, $r = 0.597$, which is substantially lower than $r = 0.940$ as calculated by Fleagle *et al.* This leaves >60% of the variance unexplained, which in turn indicates that the coefficient of variation is a much poorer predictor of the degree of sexual dimorphism than they suggest. A solid quantitative method for estimating the degree of sexual dimorphism in fossils based on the variability within a species and independent of prior sex determination of individuals still eludes us.

WALTER LEUTENEGGER

Department of Anthropology,
University of Wisconsin,
Madison, Wisconsin 53706, USA

1. Fleagle, J. G., Kay, R. F. & Simons, E. L. *Nature* **287**, 328–330 (1980).

2. Leutenegger, W. & Kelly, J. T. *Primates* **18**, 117–136 (1977).

FLEAGLE *ET AL.* REPLY—Leutenegger correctly points out that we inadvertently mixed the body size data in Table 1 of our paper¹ and therefore overestimated the correlation coefficient. In rechecking our original data, it turns out that we incorrectly transcribed the c.v. of mandibular depth for *Cebus apella* as well; the correct value is 14.5. When the

correlation coefficient for the entire sample was recalculated, using the cubed roots of body weight as suggested by Leutenegger, the correct value for r was 0.755. Although this value is lower than that previously reported, it is still highly significant ($P < 0.01$) and demonstrates a positive correlation between c.v. of mandible depth and sexual dimorphism in body weight. This correction in no way affects our major points: that the coefficient of variation for mandible depth and the amount of canine dimorphism in the Fayum anthropoids are considerably higher than that found in any of the monagamous taxa.

JOHN G. FLEAGLE

Department of Anatomical Sciences,
Health Sciences Center,
State University of New York
at Stony Brook,
Long Island, New York 11794, USA

RICHARD F. KAY

Department of Anatomy,
Duke University Medical Center,
Durham, North Carolina 27710, USA

ELWYN L. SIMONS

Duke University Center for the
Study of Primate Biology and History,
3705 Erwin Road,
Durham, North Carolina 27705, USA

1. Fleagle, J. G., Kay, R. F. & Simons, E. L. *Nature* **287**, 328–330 (1980).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

Enzyme heterozygosity and morphological variance

HANDFORD¹ recently reported a lack of association between enzyme heterozygosity and measures of morphological

variance in the Rufous-collared sparrow. He suggested that this negative case, which was in apparent contrast to the findings of myself² and Mitton³, could be explained by an inherent difference, with regard to this relationship, between homoiotherms and poikilotherms. Although Handford makes several statistical points worth consideration, he ignores the fundamental problem of statistical power in his analysis. Essentially, his data set is so small that unless we assume the putative relationship between heterozygosity and morphological variability to be extremely strong, the ability to reject the null hypothesis of equal variances will be very restricted.

If enzyme heterozygosity does exert some influence on morphological variability, our biological intuition suggests that the contribution of individual loci will be subtle. Examination of the data sets for the monarch butterfly² and the killifish³ clearly support this notion. For example, the ratio of homozygote to heterozygote variances in the monarch is only 1.17 for the entire set of loci. This does not diminish the evolutionary importance of such a relationship, but rather places realistic expectations on the magnitude of the effect. Furthermore, because we are comparing variances, the ability to make strong statistical inferences about such associations will be constrained without large sample sizes and many independent observations.

The statistical power for the variability measures and sample sizes reported by Handford² is difficult to quantify. Regardless of the drawbacks of the F -test, it can be used in a simple univariate case to illustrate the limitations of these sample sizes. For a normally distributed character, the homozygote/heterozygote F ratios associated with Handford's larger sample sizes which would be statistically significant ($P < 0.05$) are 1.8 (d.f. = 130, 22) and 1.5 (d.f. = 76, 49). These are unrealistically large ratios given the subtlety of the relationship expected and observed in other organisms. The problem of type II error does not disappear by using Handford's particular treatment or by applying the sign test. In addition, Handford uses multivariate measures of variability which further reduces the statistical power with the inclusion of each new character. Indeed, three of the five tests for matrix homogeneity in females cannot be carried out because the sample sizes are so small that there are insufficient degrees of freedom.

His cautionary statement that one nonsignificant result does not negate the relationship is true, but a nonsignificant result for a truly powerful data set would have permitted some strong inference to

be made. Obviously, the need is not for many more tests of weak statistical power, but rather for a few tests, each with sufficient power to distinguish between no relationship and one of small, yet realistic, effect.

WALTER F. EANES

*Museum of Comparative Zoology,
Harvard University,
Cambridge, Massachusetts 02138, USA*

1. Handford, P. *Nature* **286**, 261–262 (1980).
2. Eanes, W. F. *Nature* **276**, 263–264 (1978).
3. Mitton, J. B. *Nature* **273**, 661–662 (1978).

HANDFORD REPLIES—Of course, I agree with what I take to be Eanes' main point: that many more data are needed to clarify the relationship between heterozygosity and morphological variability. Indeed, my modest contribution¹ closed with an implied request that workers (and there must be many) in possession of data sets containing the necessary information (for many taxa and based on larger samples) do the necessary sums and communicate their results. In that note¹ I also made the obvious point that lack of a significant result does not demonstrate much on its own (even though nonsignificance is also evident from the sign test² used by Mitton³ in the interpretation of his results).

Regarding statistical considerations, Eanes is correct about necessary magnitudes of effects when using the *F* ratio. Comment has been made about the use of *F* ratios¹, and will not be repeated. However, I may compare my data with those of Mitton³. If his three localities are considered separately, his samples are comparable, both biologically and numerically, with mine¹. In all three samples, Mitton's data yield a result which is significant by the sign test; mine do not. The sign test is, of course, independent of the magnitude of the differences in variance and of the size of the sample used in computing the variances. Given that there are some *a priori* reasons for expecting that heterozygosity may be of different biological significance among homoiotherms and poikilotherms¹, I felt my results worth communicating.

PAUL HANDFORD

*Zoology Department,
University of Western Ontario,
London, Ontario, Canada N6A 5B7*

1. Handford, P. *Nature* **286**, 261–262 (1980).
2. Sokal, R. R. & Rohlf, F. J. *Biometry* (Freeman, San Francisco, 1969).
3. Mitton, J. B. *Nature* **273**, 661–662 (1978).

Nigral-derived muscimol circling—ipsiversive, contraversive or controversial?

REAVILL *ET AL.*'S¹ attempted reconciliation of the dispute between Waddington² and Martin and Haubrich³ raises some important issues. Briefly, the controversy centres on whether disinhibition of dopamine (DA) neurones in the substantia nigra zona compacta (SNC) contributes to the contraversive circling produced by injecting a γ -aminobutyric acid (GABA) agonist into one nigra, as this has been variously reported as being unaffected^{4–6}, attenuated^{7–11} or even exaggerated¹² by measures which reduce nigrostriatal dopaminergic activity. We, too, firmly believe that the precise location of injection of the drug within the nigra holds the key to these disparities. Thus when muscimol (40 ng) is microinjected discretely (in 0.2 μ l over 3 min) into the central region of the zonal reticulata (SNR) at all rostrocaudal levels, it evokes rapid contraversive circling which is significantly (though never completely) inhibited by pretreating the animals with haloperidol (0.1 mg per kg) or pimozide (0.25 mg per kg), intraperitoneally, or 6-hydroxydopamine (4 μ g into the SNC or medial forebrain bundle). That is, it is partially DA dependent (see refs 7–11). On the other hand, injecting the agonist into ventral or lateral aspects of the SNR, irrespective of rostrocaudal coordinates, again stimulates rats to move in contralateral circles, but in this case the rate of rotation is regularly slower and is unaffected by the above treatments. In other words it is DA independent (see refs 4–6). Consequently we believe it is the disposition of muscimol in a mediolateral or dorsoventral direction within the reticulata, and not along the rostrocaudal plane of the nigra as a whole as Reavill *et al.*¹ surmised, that determines whether the resultant circling behaviour involves dopaminergic activation.

The latter could occur by muscimol suppressing the activity of inhibitory interneurons, or of nigrothalamic fibres that send collaterals to the SNR. Our suggestions are in keeping with the known topography¹³ and electrophysiology¹⁴ of this area of the nigra. The assertion that Martin and Haubrich could only have achieved striatal DA release and contraversive circling with muscimol from the caudal nigra is therefore unfounded. Furthermore, our arbitrary division of the SNR into two longitudinal bands from which different types of contraversive circling may be elicited should not be confused with the separate, weaker depressant action of muscimol on the compacta DA neurones directly¹⁴. This will give rise to ipsiversive asymmetry that is totally abolished by neuroleptics⁵. We would agree that this type of ipsiversive

behaviour is most readily evoked from the rostral SNC. In fact, we have never observed ipsilaterally directed movements with muscimol from the caudal SNC. Thus a behavioural distinction can be made between GABAergic drugs acting on the SNC (rostrally) or the SNR (caudally), but this does not account for the above-mentioned discrepancies¹.

To dismiss as 'redundant' the significance of striatal DA release in the mechanism of nigral GABA-mediated circling could be premature, as this presupposes that all such behavioural influences are necessarily channelled back through the SNR, which has yet to be ascertained. This might be true for rotational behaviour generated artificially from the SNR by an agonist as powerful as muscimol, even in 'thalamic' rats¹⁵, but we cannot help wondering if animals bearing chronic cerebral damage can validly be used to predict the importance of events that take place in the whole animal. The elegant experiments of Grace and Bunney¹⁴ favour the opposite view by showing that the firing rates of compacta DA cells are reciprocally related to the activity of non-dopaminergic GABA-receptive cells in the neighbouring SNR. The focal ejections of GABA which raised DA cell firing severalfold were made into the central region of the SNR, the area most densely innervated by striatonigral GABA terminals¹⁶. Cells here are exquisitely sensitive to iontophoretic GABA and it is from this region that we obtained contraversive turning to muscimol which was strongly dependent on dopaminergic mechanisms. There is every reason to suppose, therefore, that these GABA-DA interactions are far from inconsequential in normal physiological conditions.

M. S. STARR

I. C. KILPATRICK

*Department of Pharmacology,
The School of Pharmacy,
29–39 Brunswick Square,
London WC1N 1AX, UK*

1. Reavill, C., Leigh, N., Jenner, P. & Marsden, C. D. *Nature* **287**, 368 (1980).
2. Waddington, J. L. *Nature* **283**, 696–697 (1980).
3. Martin, G. E. & Haubrich, D. R. *Nature* **283**, 697 (1980).
4. Oberlander, C., Dumont, C. & Boissier, J. R. *Eur. J. Pharmacol.* **43**, 389–390 (1977).
5. Reavill, C., Jenner, P., Leigh, N. & Marsden, C. D. *Neurosci. Lett.* **12**, 323–328 (1979).
6. Waddington, J. L. *Eur. J. Pharmacol.* **58**, 327–329 (1979).
7. Arnt, J. & Scheel-Krüger, J. *Psychopharmacology* **62**, 267–277 (1979).
8. James, T. A. & Starr, M. S. *Nature* **275**, 229–230 (1978).
9. Kilpatrick, I. C., Starr, M. S., Fletcher, A., James, T. A. & MacLeod, N. K. *Expl. Brain Res.* **40**, 45–54 (1978).
10. Martin, G. E., Papp, N. L. & Bacino, C. B. *Brain Res.* **155**, 297–312 (1978).
11. Olpe, H.-R., Schellenberg, H. & Koella, W. P. *Eur. J. Pharmacol.* **45**, 291–294 (1977).
12. Olanas, M. C., De Montis, G. M., Mulas, G. & Tagliamonte, A. *Eur. J. Pharmacol.* **49**, 233–241 (1978).
13. Faull, R. L. M. & Mehler, J. N. R. *Neuroscience* **3**, 989–1002 (1978).
14. Grace, A. A. & Bunney, B. S. *Eur. J. Pharmacol.* **59**, 211–218 (1979).
15. Papadopoulos, G. & Huston, J. P. *Neurosci. Lett.* **13**, 63–67 (1979).
16. Ribak, C. E., Vaughan, J. E. & Roberts, E. *Brain Res.* **192**, 413–420 (1980).

BOOK REVIEWS

Oil in the wheels of politics

Ian Smart

THOSE who have to catalogue or identify books would sometimes welcome a little more care in choosing titles. Certainly, no one should reach for this stimulating but strangely uncertain symposium in the belief that it will offer a synoptic view of the relationship between energy and security. On the supply side, coal, gas, nuclear power and "renewables" are mentioned in passing, but the emphasis is otherwise exclusively on oil. On the demand side, changes in oil and total energy consumption are duly noted, but the structure and evolution of conversion and end-use patterns are largely ignored. So are all questions of energy technology. As to security, it is understandable that most of the discussions and policy recommendations relate to the United States, albeit against the background of praiseworthy insistence that American isolation from world energy concerns is impossible. What is less acceptable is the absence of any serious reflection on what "security" really means in the energy context. Some of the groundwork is here, for instance in the occasional distinction between import dependence and supply vulnerability. In the end, however, the only definition of energy security, implicitly or by exclusion, is in the all too familiar form of immunity to a sudden disruption of flows or prices in the world oil market. The result is a book about short-term threats to oil trade, with particular reference to the United States. That is undeniably an important subject, but only a broader analysis could hope to determine its true significance.

Any volume made up of separate papers derived from a series of seminars and written by both senior luminaries and graduate students confronts its editors with problems of coherence and consistency. This is no exception. In minor matters — mis-spellings or misprints — the standard is sometimes deplorable. In more important respects, however, there is better evidence than usual of thematic continuity. Nor are the more junior contributors disgraced. Indeed, Fen Hampson on Canadian energy policy and, outstandingly, Kevin Middlebrook on US-Mexican energy relations make two of the most useful and polished contributions to the collection. Some of the problems nevertheless prove insuperable. It has to be said, for example,

Energy and Security. Edited by David A. Deese and Joseph S. Nye. Pp.489. ISBN 0-88410-640-3/0-6660-684-5. (Ballinger/Harper & Row: 1981.) \$14.50, £9.50.

that a book so narrowly focused on oil politics deserves a less superficial survey of politics in the Gulf: a chapter which might, moreover, have carried more weight if, like other sections, it had been reviewed since the outbreak of the Iraq-Iran war. In general, however, the quality of the individual contributions is sufficient, and sometimes exemplary. The intractable difficulty is rather one of reconciling what is said or, more often, implied in the different papers.

In some cases, the impression created is that apparent conflicts of interpretation have been recognized but have resisted resolution. In their summing-up of lessons for American policy, for instance, the editors (with Alvin Alm) strongly endorse the establishment and use of national oil stockpiles, but without referring to some hard questions about that policy raised earlier by Thomas Neff, in a strong paper on recent changes in the oil market. In other cases, the impression is that the participants have simply not addressed themselves to fundamental questions begged by the conjunction of their separate analyses. Two such questions stand out. One is about the relationship between governments and energy markets. And the other is about the relationship between immediate protection against oil-trade disruptions and the ultimate security of energy systems at large.

The contributors are especially uncertain about the role of government. In his opening chapter, Joseph Nye argues for more government involvement in external energy affairs, but less in domestic ones. Other authors echo that view, extolling the usefulness of market forces and casting sceptical glances at the ability of government to regulate or *a fortiori* to ration the flow of oil in an emergency, or even to shape the domestic energy economy in the longer term. Yet, time and again, recommendations for policy action — on stockpiling, pricing, taxation, revenue distribution or the promotion of alternative supply — imply a need for more, rather than less, governmental intervention. The domestic market must be

made free, so that it can transmit price signals efficiently to consumers, but the market must be managed, lest the signals it transmits, as in either the easy days of 1975-1978 or the panic of 1978-1979, reflect an inefficiently short perspective. (David Deese and Linda Miller, in discussing the West European dimension, even complicate the issue by suggesting that, in an emergency, available oil will gravitate towards "free market countries".) Nowhere is there convincing evidence that the contributors and editors have faced up to the obvious dilemma lurking in those shadows. The suspicion they leave is of having convinced themselves that government, in the abstract, must play a role of which American governments, in the real world, are incapable.

The second fundamental but unresolved issue is even more obvious. Several of the authors, led by William Hogan, apparently believe that short-term measures to prepare for disruptions of oil trade have been neglected in favour of efforts to limit oil imports or oil consumption in the longer term. In the particular case of the United States, given the fiasco of the Strategic Petroleum Reserve, it would be hard not to sympathize with them. What is lacking, however, is a considered sense of the opportunity cost of short-term oil emergency measures in relation to the longer-term management of fuel supply and of the energy economy as a whole. A country's relative security in regard to energy palpably depends on much more, in the final analysis, than the scale and pattern of its oil imports. Far larger social, political and institutional, as well as economic, questions are at stake. In this case, however, they remain unilluminated. Anyone who wants to think more deeply about the problems of coping with another Arab oil embargo or the effects on oil supply of another Middle Eastern war, especially in the United States, should read this book. But oil supply is not the only element in national, let alone global, energy security, and anyone who wants to learn about other elements of that now perennial issue will look here in vain. □

Ian Smart is an independent adviser on international energy affairs.

States of matter inside the planets

W.B. Hubbard

Interiors of the Planets. By A.H. Cook. Pp.348. ISBN 0-521-23214-7. (Cambridge University Press: 1980.) £25, \$59.50.

Interiors of the Planets, a volume in the Cambridge Planetary Science series, is intended as a textbook for advanced students. The book reads rather more like an extended review article than a text, however; yet probably no other approach is possible in this rapidly-developing field.

A number of thorough, up-to-date accounts of important topics are scattered through the book. The initial discussion of the Earth and the derivation of interior models from seismic data and the Adams-Williamson equation is lucid, as is the subsequent coverage of heat sources in the Earth and effects on the temperature gradient. The strongest parts of the book are those which deal with measurement of gravitational field parameters and with solid-body planetary motions — the use of artificial satellites, fly-bys and laser retro-reflectors to determine gravitational field parameters and precession and libration constants is considered in some detail; and the first-order theory of figures, leading to the important Radau-Darwin approximation, is given in enough detail to make clear its physical significance.

There are two chapters on equations of state — separated by chapters on the Moon and terrestrial planets — a division which seems a little curious, since similar physical principles are involved. Nevertheless, the exposition of experimental techniques for measuring equations of state and phase transitions is a helpful guide to the intricate transformations which can occur in minerals at high pressure. The chapter on metallic and molecular hydrogen, which comes later, is also fairly complete, and gives the reader a good feeling for the convergence of very different methods for deriving pressure-density relations and other important thermodynamic parameters.

The chapter on the Moon is somewhat sketchy. Lunar seismology and the evidence for a lunar core from the moment of inertia are presented in useful detail, but an adequate discussion of important work on electromagnetic sounding of the lunar interior is lacking. The account of the formation of the lunar crust tends to be unclear — and many workers in this field would doubt the assertion that the crust may have accreted separately onto the Moon, from meteorites, after formation of the mantle.

The discussion of Mars, Venus and Mercury is likewise rather uneven. Moments of inertia are dealt with well, and the author succeeds in giving a clear, simple exposition of the relation between the moment of inertia and the relative effects of crust, polymorphic phase transitions

and core. On the other hand, the effects of tidal dissipation are ignored, and the classic work of Goldreich and collaborators on spin-orbit coupling in terrestrial planets is mentioned only obliquely. Although some attention is given to the higher-order gravity fields of the Earth, Moon and Mars, a geoid map of each is needed for clarity; such a map is given for the Moon only, but without definition of units.

Similarly, the chapter on the Jovian planets has some fairly major gaps. Most discussions of the composition of this group start with the mass-radius diagram for self-gravitating spheres at zero temperature. The existence of a pronounced maximum radius for any such object at about the Jovian radius is convincing evidence for the preponderance of hydrogen in Jupiter and Saturn. Without such a diagram, the reader must take this important result on trust. In

addition, the discussion of the heat balance of Jupiter and Saturn is misleading, and it is quite impossible for the solar wind or ultraviolet radiation to contribute in a significant way to the observed imbalance.

The author makes it clear that *Interiors of the Planets* is not a cosmogonically oriented book. His general theme is the use of established physical principles to draw general conclusions about the present interior state of planets. A synthesis of the derived interior structures of planets does require cosmogonical considerations, but we do not yet know enough to carry out such a synthesis.

The book will not suffice as a sole reference to planetary interiors for an interested student, but it will be a helpful addition to his library. □

W.B. Hubbard is a Professor of Planetary Sciences at the University of Arizona.

Logic in Galilean scholarship

A. Rupert Hall

Galileo and the Art of Reasoning: Rhetorical Foundations of Logic and Scientific Method. By Maurice A. Finocchiaro. Pp.478. Hbk ISBN 90-227-1094-5; pbk ISBN 90-227-1095-3. (Reidel: 1980.) Hbk Dfl.80, \$42; pbk Dfl.40, \$21.

THIS is, as befits the series in which it is published, very much a book by a philosopher for philosophers, a book about logic rather than about Galileo. A scientist, a historian or a general reader wishing to make an initial exploration of Galileo's style as a writer and a natural philosopher would do much better to sit down with an English version of Galileo's *Dialogues on the Two Chief Systems of the World* than to read this volume. Indeed, Professor Finocchiaro's study would be impossibly hard to follow for anyone who has neither familiarity with Galileo's book, to which it is directed, nor with contemporary scholarship in history and philosophy of science.

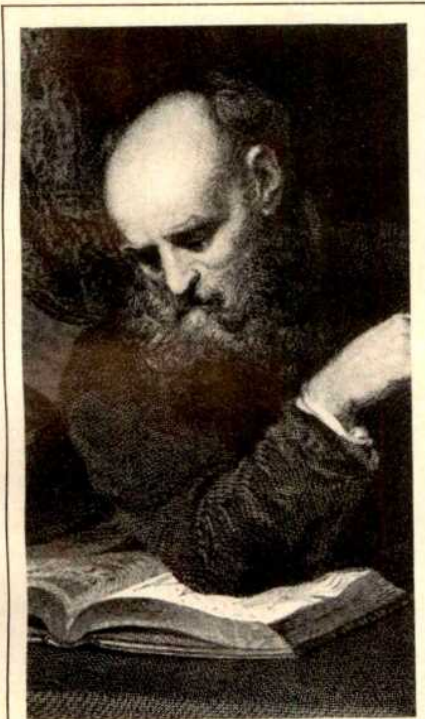
This is, in brief, a work of detailed and technical criticism, mostly of the logical structure of arguments offered either by Galileo himself or by would-be interpreters of Galileo, both historians and philosophers (and at least one great scientist — Einstein). Finocchiaro exercises singular skill and acuity in the close analysis of reasoning in order to test whether opinions and generalizations purportedly justified by it are valid or otherwise: most commonly he finds them otherwise. Thus he finds grounds for strong reservations about statements or views relating to

Galileo put forward by Alexandre Koyré, Paul Feyerabend, Maurice Clavelin, Dudley Shapere, William Shea, Stillman Drake and many others. To give one example of the elaboration involved: a quotation from a book by Koyré published 42 years ago occupies half a page, its dissection 14½ pages (admittedly of larger print). It is the fine structure of the Koyré passage that is faulted: "it would be unhistorical" the author concludes "to deny that the study of the history of science made great progress with Koyré: to turn the clock backwards is simply unthinkable".

It seems that his object is to make philosophical judgements upon the philosophical positions adopted by other writers on Galileo: Koyré's error lay in his being an apriorist rationalist; Feyerabend's in denying the rationality of Galileo's and modern science; Drake's in being an empiricist or perhaps totally aphilo-sophical, and so forth. This makes the road to a positive understanding of Galileo's reasoning long and hard. On the other hand, Finocchiaro does express plainly his sense of the merits of previous scholars, even though (in his view) all have made serious mistakes, and he has studied the *Dialogue* and the difficulty of translating it with exemplary care, providing very full analyses of both the content and the nature of the arguments used by Galileo.

There are, to me, paradoxes in Finocchiaro's book apart from the paradoxes of Feyerabend, the Zeno of out time. Why does he object to Koyré's writing that the *Dialogue* is "not a book of astronomy, nor even of physics. It is above

all a book of criticism, a work of polemic and struggle; it is at the same time a pedagogical work, and a philosophical work. . . . Why does he write of Galileo's allegedly "Platonic" hypothesis of the planets all being accelerated into their orbits by falling from a unique point in space that "the correspondence is amazingly close"? (It is not, as Newton pointed out long ago in his second letter to Richard Bentley, though Galileo seems to accept it as true.) Why is it profitable or illuminating to treat Galileo's *Dialogue* in case-study fashion as though it were a statement of contemporary science?



Mary Evans Picture Library

It is fair to end with some (abbreviated) conclusions of the author's own:

The evidence from Galileo indicates that science is primarily a method rather than a set of abstract truths; that method involves not the fixed adherence to some formal universal rules, but rather the judicious balancing of such opposites as speculation and observation, quantitative analysis and qualitative considerations. . . ; that in particular such judgment does not automatically exclude rhetorical persuasion or aesthetically expressed emotion; but that such exercise of judgment reduces ultimately to a matter of reasoning, namely the drawing of conclusions from premises and the formulation of reasons for claims.

Finocchiaro ends with the admonition that "in order to become more scientific, the study of reasoning needs a greater empirical orientation towards reasoning as it actually occurs in the world" to be found by employing a "broadly conceived historical approach" not shying away from "attempts at concrete theoretical systematizations". □

A. Rupert Hall is Emeritus Professor of the History of Science, University of London.

Essential techniques in diazotrophy

John Postgate

Methods for Evaluating Biological Nitrogen Fixation. Edited by F.J. Bergersen. Pp.702. ISBN 0-471-27759-2. (Wiley: 1980.) £38, \$114.

To do justice to a book as important as this in a short review is a frustrating task. It is a compendium of modern methods of studying diazotrophy and, as most readers will know, it is largely new (or improved) techniques which have permitted the spectacular advances of the past few decades.

First, a brief survey of its authors and scope. Dalton describes techniques for the culture of nitrogen-fixing bacteria; Bergersen discusses direct nitrogen analyses (for example Kjeldahl, $^{15}\text{N}_2$); Turner and Gibson deal with indirect assays of diazotrophy (acetylene reduction, hydrogen evolution); Gibson copes alone with methods for handling legumes in a laboratory or glasshouse; Bond and Wheeler give details for examining actinorhizal symbioses.

More biochemically, Eady comprehensively surveys methods of extracting, purifying and studying nitrogenase; Appleby and Bergersen deal with leghaemoglobins; Farnden and Robertson attempt, with considerable success, the almost impossible task of deciding what enzymes are metabolically related to nitrogenase, and tell how to handle them — they give priority to plant enzymes. Schwinghamer and Dudman then make a somewhat discursive job of "methods for identifying strains of diazotroph", straying rather often from true practicality and regrettably concerning themselves almost exclusively with rhizobia. In contrast, Cannon's terse chapter on everything you need to know to study the genetics of diazotrophs is a masterly laboratory manual in itself.

Brockwell's wide experience informs his chapter on crop and pasture legumes, which takes the reader out into the field; a comparable authority pervades Thompson's chapter on legume inoculants. Dobereiner describes the relatively new if controversial area of diazotrophic associations involving grasses and grain crops — her account of the much-criticized excised root acetylene assay, as well as other techniques, is useful and balanced. Knowles's chapter on plant communities and soils extends coverage to natural ecosystems; finally, Stewart presents a synoptic view of the cyanobacteria and their associates, detailed and down to earth on the free-living types, less practical on the symbioses.

With such broad coverage some defects are inevitable and I must mention a few. Someone ought to have warned about the hazards of "ghosts", those bacteria which scavenge atmospheric fixed N (NH_3 or

NO_x) and so simulate diazotrophs, for they are still responsible for occasional false reports of diazotrophy. Bergersen should have mentioned the use of ammonia electrodes, and also warned researchers not to smoke, sweat or permit reagent ammonia in the room (or work near stables!) when doing Kjeldahl analyses. Thompson ought to have warned of the inherent error of Petroff-Hauser chambers (they overestimate). Indeed, enumeration of diazotrophs can be tricky and it is inadequately discussed. Dobereiner considers the enumeration of some, but the topic ought surely to have been in Dalton's brief. Dalton also omitted the important problem of colonial multimorphism in putatively pure strains of rhizobia: Dr Elkan has shown that they can have different phenotypes and require special detergent media to be seen. Finally, despite its use in California, I question Farnden and Robertson's recommendation of a tritium exchange to assay uptake hydrogenases: tritium uptake is more sensitive and theoretically sounder.

No doubt other criticisms could be made, but they all become trivial in the context of the excellence of the book as a whole. Dr Bergersen is to be congratulated on choosing authors with direct, practical experience of their subjects and ability to communicate that experience to the reader. The information in this book will have a very much longer half-life than that in most of the more theoretical books and the symposia on diazotrophy which are now being published with such regularity. This one is essential for any laboratory teaching or researching in this area and even experienced research workers will benefit from dipping into it. □

John Postgate is Professor of Microbiology at the University of Sussex and Director of the Agricultural Research Council's Unit of Nitrogen Fixation.

T-cell controversies

Werner Haas

Regulatory T Lymphocytes. Edited by Benvenuto Pernis and Henry J. Vogel. Pp.449. ISBN 0-12-551860-9. (Academic: 1980.) \$47.50, £26.50.

THIS collection of papers and review articles about suppressor (T_S) and helper (T_H) T cells was written by a group of experts, mainly from the USA, who met in June 1979 at Columbia University. The majority of findings described are poorly understood and controversial; they will challenge the expert and confuse the non-expert.

Outstanding is the article by Sprent indicating that T_H interact only with B cells and macrophages expressing Ia antigens when encountered by the T cells in early differentiation. Confirmatory and contradictory experiments are discussed by Pierce and Kapp, by Mitchison and by Marrack *et al.* A more detailed analysis of regulatory T cells requires *in vitro* experiments. As yet, informative data are rare, but there are some studies in the right direction — selection in long-term cultures (Augustin *et al.*) and limiting dilution analysis of specific T_H (Zauderer *et al.*), and isolation of a T_S clone (Cantor and Gershon) are discussed here. Current experiments concerned with the serology and biochemistry of antigen-specific T cell receptors and factors are referred to in several articles and some details are described by Cramer and Krawinkel, Goodman *et al.* and Geha. Melchers and McDevitt present a clearly written, comprehensive review on the selective expression of Ia antigens by T cells. Functionally distinct human and murine T cell subsets identified by a variety of antisera and monoclonal antibodies feature in papers by Micklem *et al.*, Friedman *et al.* and Reinherz and Schlossmann; identification of subsets by red cells coated with different classes of antibodies is considered by Fauci and by

Moretta *et al.* An impressive amount of work, presented or reviewed by Fauci *et al.*, Siegal and Siegal, and Waldmann *et al.*, suggests that a variety of human diseases are associated with or caused by various forms of "disordered immunoregulation".

There is general acceptance of Jerne's notion that antigen-specific receptors must be the primary target of any specific regulatory mechanism in the immune system. This experimentally demanding issue is considered in more detail in articles on idiotype-specific T_S by Bona and Paul and T_H by Eichmann and Simon, Ig dependent T_H by Janeway *et al.*, allotype suppression by Tokuhisa *et al.* and T cell mediation of resistance to graft versus host disease by Wilson *et al.* Finally, Herzenberg and colleagues demonstrate how to express complex ideas in clear drawings.

The book will be a disappointment to readers who expect more definitive reports on the various types of T_H and T_S , H-2 restriction of T_S , inducer cells and feedback inhibition, regulation of cytotoxic T cell responses or on the various types of T cell factors inducing lymphocyte proliferation and/or differentiation. □

Werner Haas is at the Basel Institute for Immunology.

after-discharges in the same system. There is, however, a danger in generalizing from epileptics to normals since there must be something abnormal in the epileptic brain to cause the discharge in the first place.

In the absence of interesting new results to report, most of the contributors resort to the banal. It is no surprise that the brain waves produced in transcendental meditation resemble those in sleep or that they cannot be distinguished from those occurring when listening to soft music or during progressive relaxation. Nor is it particularly revealing that Italians asked to endure high levels of shock are sensible enough to call off the experiment sooner than do Yankee Protestants.

Much of the theorizing is at the same level of banality. C. E. Izard cites with approval W. Charlesworth's opinion, doubtless reached after years of wrestling with the problem, that "surprise is a misexpected object or event", whilst Karl Pribram, in one of the few intelligible sentences in his chapter, enlightens the ignorant reader with the news that "Science pursues knowledge by observation and experiment".

Most of the book's theorizing is so nonsensical that it is impossible to describe. Its level may be gauged by the following quotation from the chapter by R. J. Davidson, one of the editors:

The term 'unconscious' is here being used . . . (1) to refer to mental structures . . . which contour and transform both exogenous (from the environment) as well as endogenous (from within the person . . .) information, and (2) to refer to processing activity, action and/or information transmission in the nervous system which proceeds in the absence of conscious awareness . . . The first usage makes reference to internal mental (i.e. biocognitive) structures . . . The second usage is concerned with observable behaviour (either directly or indirectly) of which the subject is unaware.

One cannot evaluate Dr Davidson's claim to have uncovered two kinds of unconscious process, since he is clearly unable to put the difference into words: nor does he explain what is meant by "observable behaviour of which the subject is unaware". In an equally silly chapter, Pribram argues that "feedforward" operations require consciousness: the reader who wants to know what feedforward is learns that it is "feedbacks joined into parallel processes".

Much of the book — and all of its theoretical sections — is a farrago of pretentious rubbish; it is atrociously written, with the usual attempt to make commonplace ideas sound important by the introduction of neologisms. In short, the book is a disgrace to the publishers, the editors and the contributors. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.

Unconscious pretension in psychology

Stuart Sutherland

The Psychobiology of Consciousness. Edited by Julian M. Davidson and Richard J. Davidson. Pp.489. ISBN 0-306-40138-X. (Plenum: 1980.) \$32.50, £20.48.

ONE of the most significant changes in psychology over the past two decades is that consciousness has become respectable. Several separate developments have brought it back to fashion, including the realization that behaviourism can offer no explanation for many actions, the development of computer simulation as a method for modelling complex thought processes, the increasing interest in Western society in attaining "altered states of consciousness" either through psychotropic drugs or by meditative practices, and the discovery of many different transmitter substances that are involved in the regulation of mood. It is with the latter two themes that *The Psychobiology of Consciousness* is mainly concerned.

The field appears to be in such chaos that the attempt to think about it has reduced the book's 15 contributors to incoherence. Their strenuous efforts to resuscitate the shade of Titchener by recording and classifying states of consciousness, both normal and abnormal, have met with no success. The few interesting new findings

they present have little unity. Typical examples of such isolated discoveries are that since large quantities of marijuana usually have little effect on novices, new users of the drug have to learn how to become "stoned", and that there is a tendency to look to the right (which may mean that the left hemisphere is being used) when thinking verbally, and to the left when working out spatial problems.

Perhaps of more significance are some recent studies of temporal lobe epilepsy. Synchronous spiking of nerve firing within the temporal lobe can be accompanied by feelings of ecstasy, religious awe and extreme kindness, and it is usually associated both with a lack of interest in sex and with the absence of any form of addiction. Treatment of the spike focus to eliminate the hypersynchrony often alters the epileptic's personality for the worse. His irritability increases, and he may become unhappy: he regains an interest in sex, but usually loses his religious preoccupations. Although A.J. Mandel does not go so far as to allege that St Paul had a seizure on the road to Damascus, he does argue that ecstasy is normally accompanied by hypersynchronous discharge in the limbic system excluding the amygdala and that the prolonged state of saintly kindliness that ensues depends upon

BOOKS RECEIVED

Physics

- DIEULESAINT, E and ROYER, D Elastic Waves in Solids Applications to Signal Processing Pp 511 ISBN 0-471-27836-X (Wiley-Interscience 1980) £19 50, \$59
- HAGEDORN, P Non-Linear Oscillations Pp 289 ISBN 0-19-856142-3 (Clarendon Press/Oxford University Press 1981) £27 50
- HÖCHSTRASSER, R, KAISER, W and SHANK, C V (eds) Picosecond Phenomena II Springer Series in Chemical Physics, Vol 14 Proceedings of the 2nd International Conference, Cape Cod, Massachusetts, June 1980 Pp 382 ISBN 3-540-10403-8 (Springer-Verlag 1980) DM 69, \$40 80
- LAITHWAITE, E R and FRERIS, L L Electric Energy Its Generation, Transmission and Use Pp 365 Flexi ISBN 0-07-084109-8 McGraw-Hill 1980) £9 95
- LANDAU, L D and LIFSHITZ, E M Course of Theoretical Physics, Vol 9 Statistical Physics Part 2 Pp 350 Flexi ISBN 0-08-023072-5 (Pergamon 1979) \$20, £9
- MANN, H K (guest ed) Radiation Physics and Chemistry Radiation Sterilisation of Plastic Medical Devices Pp 124 Flexi ISBN 0-08-025067-X (Pergamon 1980) £7 50
- MARLOW, W H (ed) Aerosol Microphysics I Particle Interaction Topics in Current Physics, Vol 16 Pp 160 ISBN 3-540-09866-6 (Springer-Verlag 1980) DM 54, \$31 90
- MARTON, L and MARTON, C (eds) Advances in Electronics and Electron Physics, Vol 51 Pp 455 ISBN 0-12-014651 (Academic 1980) \$49 50
- MEZEI, F (ed) Neutron Spin Echo Proceedings of a Laue-Langevin Institut Workshop, Grenoble, October 1979 Lecture Notes in Physics, Vol 128 Pp 253 Flexi ISBN 3-540-10004-0 (Springer-Verlag 1980) DM 33, \$19 50
- MOSS, T S (executive ed), KNEÜBUHL, F (guest ed) CIRP 2 2nd International Conference on Infra-red Physics Papers presented at the Conference, Zurich, March 1979 Pp 506 Flexi ISBN 0-08-025055-6 (Pergamon 1980) £12 50
- NAKAJIMA, S, TOYOZAWA, Y and ABE, R The Physics of Elementary Excitations Springer Series in Solid-State Sciences, Vol 12 Pp 332 ISBN 3-540-09921-2 (Springer-Verlag 1980) DM 85, \$50 20
- NESBET, R K Variational Methods in Electro-Atom Scattering Theory Pp 228 ISBN 0-306-40413-3 (Plenum 1980) \$32 50
- OSTROVSKY, Yu I, BUTUSOV, M M and OSTROVSKAYA, G V Interferometry by Holography Springer Series in Optical Sciences, Vol 20 Pp 330 ISBN 3-540-09886-0 (Springer-Verlag 1980) DM 89, \$52 60
- PANKOVE, J I (ed) Display Devices Topics in Applied Physics, Vol 40 Pp 252 ISBN 3-540-09868-2 (Springer-Verlag 1980) DM 89, \$52 60
- PAULI, W General Principles of Quantum Mechanics Pp 212 Flexi ISBN 3-540-09842-9 (Springer-Verlag 1980) DM 29, \$17 20
- RUDDEN, M N and WILSON, J Elements of Solid State Physics Pp 186 Hbk ISBN 0-471-27750-9, pbk ISBN 0-471-27749-5 (Wiley 1980) Hbk, np, pbk £5 95
- SANCHEZ-PALENCIA, E Non-Homogeneous Media and Vibration Theory Lecture Notes in Physics, Vol 127 Pp 398 Flexi ISBN 3-540-10000-8 (Springer-Verlag 1980) DM 41, \$24 20
- SCHECHTER, M Operator Methods in Quantum Mechanics Pp 324 ISBN 0-444-00410-6 (Elsevier/North-Holland 1981) \$32 50 (US and Canada), \$44 (rest of world)
- SCHLENKER, M *et al* (eds) Imaging Processes and Coherence in Physics Proceedings of a workshop held at Les Houches, France, March 1979 Lecture Notes in Physics, Vol 112 Pp 577 Flexi ISBN 3-540-09727-9 (Springer-Verlag 1980) DM 59, \$34 90
- SHANKAR, R Principles of Quantum Mechanics Pp 612 ISBN 0-306-40397-8 (Plenum 1980) \$29 50
- SILINSH, E A Organic Molecular Crystals Their Electronic States Springer Series in Solid-State Sciences, Vol 16 Pp 389 ISBN 3-540-10053-9 (Springer-Verlag 1980) DM 92, \$54 10
- SPRINGFORD, M (ed) Electrons at the Fermi Surface Pp 541 ISBN 0-521-22337-7 (Cambridge University Press 1980) £35
- SUBBARAO, E C (ed) Solid Electrolytes and Their Applications Pp 298 ISBN 0-306-40389-7 (Plenum 1980) \$35
- SUENAGA, M and CLARK, A F Filamentary A15 Superconductors Proceedings of the Conference Brookhaven National Laboratory, May 1980 Pp 368 ISBN 0-306-40622-5 (Plenum 1980) \$45
- SUPPES, P (ed) Studies in the Foundation of Quantum Mechanics Pp 265 ISBN 0-917586-12-3 (Philosophy of Science Association, East Lansing, Michigan 1980) \$6
- THOMPSON, G H B Physics of Semiconductor Laser Devices Pp 549 ISBN 0-471-27685-5 (Wiley 1980) £22 75
- TIMMERHAUS, K D and SNYDER, H A Advances in Cryogenic Engineering, Vol 25 Proceedings of the 1979 Conference, held at the University of Wisconsin, Madison, August 1979 Pp 852 ISBN 0-306-40504-0 (Plenum 1980) Np
- Van de HULST, H C Multiple Light Scattering Tables, Formulas, and Applications, Vol 1 Pp 299 ISBN 0-12-710701-0 (Academic 1980) \$38 50
- VINNICHENKO, N K *et al* Turbulence in the Free Atmosphere 2nd Edn Pp 310 ISBN 0-306-10959-X (Consultants Bureau, New York and London 1980) \$49 50
- VINOGRADOV, G V and MALKIN, A Ya Rheology of Polymers Viscoelasticity and Flow of Polymers Pp 467 ISBN 3-540-09778-3 (Springer-Verlag 1980) DM 98, \$57 90

WANG, K Y Acoustical Imaging, Vol 9 Visualization and Characterisation Proceedings of the 9th International Symposium on Acoustical Imaging, held in Houston, Texas, December 1979 Pp 842 ISBN 0-306-40477-X (Plenum 1980) \$75

WESSON, P (ed) Gravity, Particles, and Astrophysics A Review of Modern Theories of Gravity and G-Variability, and Their Relation to Elementary Particle Physics and Astrophysics Astrophysics and Space Science Library, Vol 79 Pp 276 ISBN 90-277-1038-X (Reidel 1980) Dfl 65, \$34

Chemistry

- BRITISH PHARMACOPOEIA 1980 Infra-red Reference Spectra Published on the Recommendation of the Medicines Commission Pursuant to the Medicines Act 1968 Pp 332 ISBN 0-11-320303-9 (HMSO, London 1980) Np
- BURDETT, J K Molecular Shapes Theoretical Models of Inorganic Stereochemistry Pp 287 ISBN 0-471-07860-3 (Wiley-Interscience 1980) £13 75, \$37
- CANNON, R D Electron Transfer Reactions Pp 351 ISBN 0-408-10646-8 (Butterworths 1980) £32
- CHERRY, B W Polymer Surfaces Pp 161 Hbk ISBN 0-521-23082-9, pbk ISBN 0-521-29792-3 (Cambridge University Press 1981) Hbk £14, pbk £6 95
- DUNITZ, J D *et al* (eds) Luminescence and Energy Transfer Vol 42, Structure and Bonding Pp 133 ISBN 3-540-10395-3 (Springer-Verlag 1980) DM 68, \$40 20
- HALL, C Polymer Materials, an Introduction for Technologists and Scientists Pp 198 Hbk ISBN 0-333-28907-2, pbk ISBN 0-333-28908-0 (Macmillan Press, London and New York 1981) Hbk £14, pbk £6 50
- HELFRICH, W and HEPPKE, G (eds) Liquid Crystals of One- and Two-Dimensional Order Springer Series in Chemical Physics, Vol 11, Proceedings of the Conference held at Garmisch-Partenkirchen, Germany, January 1980 Pp 416 ISBN 3-540-10399-6 (Springer-Verlag 1980) DM 69, \$40 80
- JENKINS, A D (ed) Progress in Polymer Science, Vol 6 Pp 257 ISBN 0-08-020335-3 (Pergamon 1980) £35, \$81
- MASTERS, C Homogeneous Transition-Metal Catalysis A Gentle Art Pp 277 Hbk ISBN 0-412-22110-1, pbk ISBN 0-412-22120-9 (Chapman & Hall 1981) Hbk £20, pbk £9 50
- MOLIN, Y N, SALIKHOV, K M and ZAMARAKEV, K I Spin Exchange Principles and Applications in Chemistry and Biology Springer Series in Chemical Physics, Vol 8 Pp 242 ISBN 3-540-10095-4 (Springer-Verlag 1980) DM 73, \$43 10
- PAMPLIN, B R Crystal Growth 2nd Edn International Series on the Science of the Solid State, Vol 16 Pp 609 ISBN 0-08-026489-1 (Pergamon 1980) £60, \$140
- TINKER, P B (ed) Soils and Agriculture Critical Reports on Applied Chemistry, Vol 2 Pp 151 Flexi ISBN 0-632-00722-2 (Blackwell Scientific 1980) £8
- VARMUZA, K Lecture Notes in Chemistry Vol 211, Pattern Recognition in Chemistry Pp 217 Flexi ISBN 3-540-10273-6 (Springer-Verlag 1980) DM 35 60, \$21
- WILLIS, R F (ed) Vibrational Spectroscopy of Adsorbates Springer Series in Chemical Physics, Vol 15 Pp 184 ISBN 3-540-10429-1 (Springer-Verlag 1980) DM 52, \$30 70
- WITTCOFF, H A and REUBEN, B G Industrial Organic Chemicals in Perspective Part Two Technology, Formulation and Use Pp 502 ISBN 0-471-05780-0 (Wiley-Interscience 1980) £24, \$55

Technology

- ONOE, M, PRESTON, J Jr and ROSENFELD, A Real-Time Medical Image Processing Based on Proceedings of the Japan-United States Seminar, held in Tokyo, October 1978 Pp 244 ISBN 0-306-40551-2 (Plenum 1980) \$35
- PENNER, S S and ICERMAN, L Energy Vol I, Demands, Resources, Impact, Technology, and Policy 2nd Edn One of a Three-Vol Set of Lecture Notes Pp 505 ISBN 0-201-05572-4 (Addison-Wesley 1980) Hbk \$28 50, pbk \$16 50

Earth Sciences

- ANDERSON, J G C and OWEN, T R The Structure of the British Isles 2nd Edn Pp 251 Hbk ISBN 0-08-023998-6, flexi ISBN 0-08-023997-8 (Pergamon 1980) Hbk £12 50, \$28, flexi £5 95, \$13 50
- CHILINGARIAN, G V and VORABUTR, P Drilling and Drilling Fluids Developments in Petroleum Science, 11 Pp 767 ISBN 0-444-41867-9 (Elsevier Scientific 1980) \$136 50, Dfl 280
- DECKER, R and DECKER, B Volcanoes Pp 244 Hbk ISBN 0-7167-1241-5, pbk ISBN 0-7167-1242-3 (W H Freeman 1981) Hbk \$17 95, pbk \$8 95
- FALKOWSKI, P G (ed) Primary Productivity in the Sea Environmental Science Research, Vol 19 Pp 531 ISBN 0-306-40623-3 (Plenum 1980) \$49 50
- LANGGUTH, H R and VOIGT, R Hydrogeologische Methoden Pp 486 Flexi ISBN 3-540-10174-8 (Springer-Verlag 1980) DM 45, \$26 60
- NRIAGU, J O (ed) Nickel in the Environment Pp 833 ISBN 0-471-05885-8 (Wiley-Interscience 1980) £34 75, \$81 25
- PHILLIPS, W R and GRIFFEN, D T Optical Mineralogy The Nonopaque Minerals Pp 677 ISBN 0-7167-1129-X (W H Freeman 1980) \$39 95

POTTER, P E, MAYNARD, J B and PRYOR, W A Sedimentology of Shale Study Guide and Reference Source Pp 306 ISBN 3-540-90430-1 (Springer-Verlag 1980) DM 59 90, \$32 80

PYLE, G F (ed) New Directions in Medical Geography Medical Geography Papers from the 75th Anniversary Meeting of the Association of American Geographers, Philadelphia, Pennsylvania, April 1979 Pp 298 ISBN 0-08-025817-4 (Pergamon 1980) £5 75, \$13

RINEHART, J S Geysers and Geothermal Energy Pp 223 ISBN 3-540-90489-1 (Springer-Verlag 1980) DM 38, \$22 50

Biological Sciences

BANNISTER, J V and HILL, H A O (eds) Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase Developments in Biochemistry, Vol 11A Proceedings of the Federation of European Biochemical Societies Symposium No 62 Pp 414 ISBN 0-444-00442-4 (Elsevier/North-Holland 1980) \$45, Dfl 92 50

HUNTER, R H F Physiology and Technology of Reproduction in Female Domestic Animals Pp 393 ISBN 0-12-361950-5 (Academic 1980) £25 (UK only), \$57 50

LARSEN, J A The Boreal Ecosystem Physiological Ecology Pp 500 ISBN 0-12-436880-8 (Academic 1980) Np

LEE, R E Physiology Pp 478 Hbk ISBN 0-521-22530-2, pbk ISBN 0-521-29541-6 (Cambridge University Press 1980) Hbk £27 50, pbk £9 95

LOCKE, M and SMITH, D S (eds) Insect Biology in the Future "VBW 80" Pp 977 ISBN 0-12-454340-5 (Academic 1980) \$50

MARSHALL, P T and HUGHES, G M Physiology of Mammals and Other Vertebrates 2nd Edn Pp 343 Hbk ISBN 0-521-22633, pbk ISBN 0-521-29586-6 (Cambridge University Press 1980) Hbk £20, pbk £7 95

MCGINNIS, M R Laboratory Handbook of Medical Mycology Pp 661 ISBN 0-12-482850-7 (Academic 1980) \$55

SELIGMAN, M and HITZIG, W H (eds) Primary Immunodeficiencies Proceedings of the International Symposium held in Abbaye de Royaumont, July 1980 Infirm Symposium, No 16 Pp 577 ISBN 0-444-80296-7 (Elsevier/North-Holland Biomedical 1980) \$83, Dfl 170

STEYN, P S (ed) The Biosynthesis of Mycotoxins A Study in Secondary Metabolism Pp 432 ISBN 0-12-670650-6 (Academic 1980) \$44

STODDART, D M (ed) Olfaction in Mammals Symposia of the Zoological Society of London, No 45 Pp 364 ISBN 0-12-613345-X (Academic 1980) £27 40 (UK only), \$63 50

TESTA, F Leaves Pp 28 ISBN 0-905478-74-6 (Andersen Press London, 1980) £3 50

THOMAS, G, PODESTA, E J and GORDON, J (eds) Protein Phosphorylation and Bio-Regulation Friedrich-Miescher Institut and EMBO Workshop, Basel, December 1979 Pp 232 ISBN 3-8055-1168-X (Karger 1980) SwFr 82, DM 98, \$49 25

THORPE, J E (ed) Salmon Ranching Pp 441 ISBN 0-12-690660-2 (Academic 1980) £32 40, \$78

WEINSTEIN, L and FIELDS, B N (eds) Seminars in Infectious Disease, Vol III Pp 323 ISBN 0-913258-69-5 (Thieme-Stratton, New York 1980) \$32 50

WINDLE, W F (ed) The Spinal Cord and its Reaction to Traumatic Injury Anatomy-Physiology-Pharmacology-Therapeutics Modern Pharmacology-Toxicology Pp 384 ISBN 0-8247-6688-1 (Marcel Dekker 1980) \$45 Price is 20% higher outside the U S and Canada

Applied Biological Sciences

BASTENIE, P A, BONNYNS, M and VANHAELST L Recent Progress in Diagnosis and Treatment of Hypothyroid Conditions Proceedings of the XVth International Congress of Therapeutics, Brussels, September 1979 Pp 155 ISBN 90-219-0448-9 (Excerpta Medica 1980) Np

BOURNE, G H, DANIELLI, J F and JEON, K W (eds) International Review of Cytology, Supplement 11A Perspectives in Plant Cell and Tissue Culture Pp 253 ISBN 0-12-364371-6 (Academic 1980) Np

BOLRNE, G H, DANIELLI, J F and JEON, K W (eds) International Review of Cytology, Supplement 11B Perspectives in Plant Cell and Tissue Culture Pp 257 ISBN 0-12-364372-4 (Academic 1980) \$29 50

CANTOR, C R and SCHIMMEL, P R Biophysical Chemistry, Part III The Behaviour of Biological Macromolecules Pp 524 Hbk ISBN 0-7167-1191-5, pbk ISBN 0-7167-1192-3 (W H Freeman 1980) Hbk £22 40, pbk £11 90

COHEN, E P and KÖHLER, H (eds) Membranes, Receptors, and the Immune Response 80 years after Ehrlich's Side Chain Theory Progress in Clinical and Biological Research, Vol 42 Pp 404 ISBN 0-8451-0042-4 (Alan R Liss, New York 1980) \$34 Available in Europe and the Middle East from European Book Service, Flevolaan, Holland, Dfl 102

CROFT, L R Handbook of Protein Sequence Analysis A Compilation of Amino Acid Sequences of Proteins with an Introduction to the Methodology 2nd Edn Pp 628 ISBN 0-471-27703-7 (Wiley 1980) £38

CROFT, L R Introduction to Protein Sequence Analysis Pp 157 Flexi ISBN 0-471-27710-X (Wiley 1980) £4 95

DAVIS, W, HARRAP, K R and STATHOPOULOS, G (eds) Human Cancer Its Characterization and Treatment Advances in Tumour Prevention Detection and Characterization, Vol 5 Proceedings of the 8th International Symposium on the Biological Characterization of Human Tumours Pp 420 ISBN 90-219-0458-6 (Excerpta Medica 1980) \$90 25, Dfl 185

DAVISON, A N and CUZNER, M L (eds) The Suppression of Allergic Encephalomyelitis and Multiple Sclerosis Based on the Proceedings of the Liversidge Symposium, held in October 1979 Pp 292 ISBN 0-12-206660-X (Academic 1980) £14 80, \$34 50

ELLIS, G P and WEST, G B Progress in Medicinal Chemistry, 17 Pp 280 ISBN 0-7204-0669-2 (Elsevier/North-Holland 1980) \$63 50, Dfl 130

FRIEDLANDER, M (ed) Immunological Approaches to Embryonic Development and Differentiation, Part II Current Topics in Developmental Biology, Vol 14 Pp 381 ISBN 0-12-153114-7 (Academic 1980) \$39 50

GREEP, R O (ed) Recent Progress in Hormone Research Proceedings of the 1979 Laurentian Hormone Conference, Vol 36 Pp 629 ISBN 0-12-57136-0 (Academic 1980) \$54

General

DORF, R C Modern Control Systems 3rd Edn Pp 493 Flexi ISBN 0-201-03147-7 (Addison-Wesley 1980) £8 95

DUNCAN, W R (ed) Soviet Policy in the Third World Pp 322 ISBN 0-08-025125-0 (Pergamon 1980) £17 50

ELKINGTON, J The Ecology of Tomorrow's World-Industry's Environment Pp 311 ISBN 08227-260-X (Associated Business Press, London 1980) £12

FEINBERG, G, SUNYAR, A W and WENESER, J (eds) A Festschrift for Maurice Goldhaber Transactions of the New York Academy of Sciences Series 11, Vol 40 Pp 293 Flexi ISBN 0-89766-086-2 (New York Academy of Sciences 1980) \$25

GIERYN, T F (ed) Science and Social Structure A Festschrift for Robert K Merton Transactions of the New York Academy of Sciences Series II, Vol 39 Pp 173 Flexi ISBN 0-89766-043-9 (New York Academy of Sciences 1980) \$15

GRASSE, P P L'Homme en Accusation De la Biologie à la Politique Pp 354 Flexi ISBN 2-226-01054-8 (Albin Michel, Paris 1980) Np

HAKES, D T The Development of Metalinguistic Abilities in Children Springer Series in Language and Communication, Vol 9 Pp 119 ISBN 3-540-10295-7 (Springer-Verlag 1980) DM 38 50, \$22 80

HAMMOND, J R (ed) H G Wells Interviews and Recollections Pp 121 ISBN 0-333-27416-4 (Macmillan, London 1980) £12

HAWRYLYSHYN, B Road Maps to the Future Pp 193 Hbk ISBN 0-08-026115-9, flexi ISBN 0-08-026114-0 (Pergamon 1980) Hbk £12 75, \$30, flexi £4 50, \$10 50

HINICH, M J and STAELIN, R Consumer Protection Legislation and the U S Food Industry Pp 134 ISBN 0-08-025093-9 (Pergamon 1980) £8

HOLDGATE, M W A Perspective of Environmental Pollution Pp 278 Hbk ISBN 0-521-22197-8, pbk ISBN 0-521-29972-1 (Cambridge University press 1980) Hbk np, pbk £5 95

IJIMA, N (ed) Pollution Japan Historical Chronology Pp 401 ISBN 0-08-02642-2 (Pergamon 1980) \$60, £26 50

JASPERS, K Die Idee der Universität Reprint der Erstauflage Berlin 1946-ergänzt um ein Geleitwort von A Laufs, Universität Heidelberg Pp 132 Flexi ISBN 3-540-10071-7 (Springer-Verlag 1980) DM 28, \$16 60

JERRARD, H G and D McNEILL, D B A Dictionary of Scientific Units (including dimensionless numbers and scales) 4th Edn Pp 212 Hbk ISBN 0-412-22360-0, pbk ISBN 0-412-22370-8 (Chapman and Hall 1980) Hbk £10, pbk £5 95

JUSCZYK, P W and KLEIN, R M (eds) The Nature of Thought Essays in Honour of D O Hebb Pp 256 ISBN 0-89859-034-5 (Lawrence Erlbaum Associates, Hillsdale, New Jersey 1980) \$24 95

KASSAM, S A and THOMAS, J B (eds) Nonparametric Detection Theory and Applications Benchmark Papers in Electrical Engineering and Computer Science, Vol 23 Pp 347 ISBN 0-87933-359-6 (Academic 1980) \$50

KAVANAU, J L Symmetry An Analytical Treatment Pp 656 Flexi ISBN 0-937292-001 (Science Software Systems, Los Angeles, California 1980) Np

KELLY, P J and SCHAEFFER, G (eds) Biological Education for Community Development Pp 191 Flexi ISBN 0-85066-214-1 (Taylor & Francis, London 1980) £9 50

LASS, N J (ed) Speech and Language Advances in Basic Research and Practice, Vol 3 Pp 311 ISBN 0-12-608603-6 (Academic 1980) \$29

LASZLO, E and KURTZMAN, J (eds) Western Europe and the New International Economic Order Pp 129 ISBN 0-08-025114-5 (Pergamon 1980) £7 50

LEVIEN, R E (Director) IIASA in 1979 Transformation, Continuity, and Progress Pp 153 (International Institute for Applied Systems Analysis, Austria 1980) Flexi Np

LORD, F M Applications of Item Response Theory to Practical Testing Problems Pp 274 ISBN 0-89859-006-X (Lawrence Erlbaum Associates, Hillsdale, New Jersey 1980) \$29 95

MARGERISON, T, WALLANGE, M and HALLENSTEIN, D The Superpoison 1976-1978 Pp 236 ISBN 0-333-22797-2 (Macmillan, London 1981) £7 95

MARSH, W D Economics of Electric Utility Power Generation Pp 190 ISBN 0-19-856130-X (Clarendon Press/Oxford University Press 1980) £20

MATURANA, H R and VARELA, F J Autopoiesis and Cognition. The Realization of the Living Pp 140 Hbk ISBN 0-277-1015-5, pbk ISBN 0-277-1016-3 (Reidel 1980) Hbk Dfl 55, \$28 95, pbk Dfl 27 50, \$14 50

MAXWELL, K E Environment of Life 3rd Edn Pp 614 Flexi ISBN 0-8185-0355-6 (Brooks/Cole, Monterey, California 1980) \$17 95

MITCHELL, J (ed) The Curious Naturalist Pp 204 Hbk ISBN 0-13-195412-1, pbk ISBN 0-13-195404-0 (Prentice-Hall 1980) Hbk \$15 95, pbk £7 95

MOORE, C B (ed) Chemical and Biochemical Applications of Lasers, Vol V Pp 281 ISBN 0-12-505405-X (Academic 1980) \$23

MORAVCSIK, E A and WIRTH, J R (eds) Syntax and Semantics Vol 13, Current Approaches to Syntax Pp 403 ISBN 0-12-613513-4 (Academic 1980) \$32

23 April 1981

Tenth birthday for the environment

1982 will be a red-letter year in the history of the United Nations Environmental Programme, for then the newest of the United Nations agencies is planning to celebrate the tenth anniversary of its foundation, at the United Nations Conference on the Environment in Stockholm. The meeting of the agency's governing council planned to open on 13 May in Nairobi will be asked formally to approve the plans that have been laid to celebrate the occasion. There is, for example, to be a "session of a special character", a week-long occasion for mutual congratulation and exhortation, sandwiched into the normal proceedings in 1982. For this jamboree, nations which are not members of UNEP itself will be invited to attend, the only necessary qualification is that they should be members of some other United Nations agency or of the International Atomic Energy Agency. Those arriving in Nairobi in 1982 will be provided with a "60-70 page report" by the executive director, dealing with past and future environmental challenges, together with the annual document in which the agency spells out its estimate of the then "state of the environment". The meeting of the governing council next month will be asked to approve an outline of these documents and also a budget for the "session of a special character" which amounts to \$442,000, including \$163,000 by way of fees to the consultants who will be asked to prepare the special documents. It will be interesting to see how many of the delegates in May consider that, at something like \$2,500 a page, the cost may be excessive.

Compared with some other United Nations agencies, UNEP is neither ostentatiously extravagant nor obviously bigger than it needs to be. The agency rubs along on a mere \$5 million a year, spending perhaps twice as much on specific programmes to which member states contribute separately and voluntarily. Plainly it is not living off the hog, while it is not outrageous, at least by the standards set by other international organizations, that it should spend the best part of \$500,000 on a self-congratulatory conference a year from now. The more important question, which the delegates to Nairobi should ask themselves in May, is whether they should not spend their "session of a special character" the following year considering whether their agency is pointed in the right direction.

The most serious danger, clearly apparent in Stockholm in 1972, was that the agency might be allowed to tackle the variety of insoluble problems on which the founding fathers had set their hearts. That, fortunately, has been avoided, chiefly because governments were unwilling to provide the necessary funds. Since then, the agency has made its way in the world with a number of good works of a more unspectacular kind. It has helped to bring the convention on Pollution in the Mediterranean into being — and in the past few weeks has been seeking to do the same for the Caribbean (see p 622). It has competently taken over from the headquarters in New York the management of the United Nations committee on nuclear radiation.

Only in two respects can it be seriously faulted. Over the years, the agency has taken the view that a large part of its objective should be to "raise the consciousness" of member governments on environmental matters, which is entirely proper. But in doing so, the agency has taken to using the rhetoric of the extremists in the environmental movement. A simple booklet, published the other day and consisting of a reprint of the Stockholm Declaration and related documents, is loosely and unnecessarily called "In defence of the Earth". And, in attempting to draw attention to environmental problems, the agency has often fallen

into the trap of dramatizing them. A session of a special character in 1982 might usefully try to bring the agency back to earth in these respects. At the same time, UNEP could usefully be pointed in the direction of tasks that need to be carried out, and which are squarely within the terms of reference with which it was blessed at Stockholm. All member states, but developing countries in particular, are much in need of some way of ordering priorities in the commitment of resources to the improvement of the environment. UNEP could usefully shoulder these simple tasks, and earn itself a reputation in the process.

Going private

Under what circumstances can a research laboratory established under the wing of a protective government be legitimately turned into a private enterprise? Since its election nearly two years ago, the British government has been doing what it can to sell off assets which it has accumulated, sometimes by accident. A large block of shares in British Petroleum, the oil company, has already been disposed of, as well as 49 per cent of the shares in British Aerospace, a conglomerate of miscellaneous manufacturing concerns. In due course, the government also plans to sell off enterprises such as Cable and Wireless, the publicly owned international telecommunications concern. The chief objective of these sales of public assets is to raise cash that can be used to span the annual gap between government expenditure and revenue — and it is an open secret that the government is disappointed that "privatization" (as the jargon has it) has so far accomplished so little. Partly as a consequence, but also because the government is rightly persuaded that the strength of the civil service should be reduced, attention has now turned to the conversion of public laboratories into private enterprises of various kinds. The immediate danger is that the government may not be able to distinguish between the shedding of a laboratory and the shedding to the stock market of the public interest in a block of shares in some enterprise whose management is otherwise unchanged.

Several public laboratories are in various stages of being turned out into the private sector. The simplest case is that of the Radiochemical Centre at Amersham in Buckinghamshire, which supplies radioactive isotopes and chemicals containing them to an impressive array of customers throughout the world. Technically, the centre is constituted as a limited company, the shares in which are held by the UK Atomic Energy Authority, acting as a kind of public trustee. The 1,500 or so employees are not civil servants in the strict sense. The enterprise makes a tidy profit (nearly £1.5 million in 1979-1980), and has a marketable record of growth over the past decade. Although potential purchasers will be disappointed that the profit is merely 3.5 per cent of the turnover, there is every chance that this sound business could be sold as a going concern without interfering in the short run with the management of the laboratory. The government's chief concern, when it invites tenders for the centre a year or so from now, is likely simply to be that the Radiochemical Centre is not snapped up by some unwelcome bidder. Privatization will be found to have been painless.

The other laboratories on whose disposal the government has set its heart will not be as simply turned out into the cold. The chief of them is the Hydraulics Research Station, which has over several

years acquired an enviable reputation for model studies of hydraulic projects, many of them overseas (see *Nature* 9 April, p 438). Over the years, successive governments have expected different things of the laboratory. Sometimes the laboratory has acted as a hidden subsidy of the export trade. Now, with the doctrine that those who call the tune should pay the piper sweeping through Whitehall, it is natural enough that this laboratory should find itself one of the first on the edge of unfamiliar and apparently hostile territory. The laboratory will not, however, be the last to find itself in this position.

The most serious doubt about the government's intentions concerns the prospectus which the staff at the laboratory is being offered. The assumption seems to be that if the laboratory were established as a private company, it would be able to make its way in the world by selling its services to those who use them now. The most obvious difficulty is that the fees which have been charged to, say, construction companies looking for technical appraisals of proposed designs have not fully covered the overhead cost, and especially the cost of long-term research. Those working at the laboratory, and the Institute of Professional Civil Servants which represents the scientists among them, are naturally acutely aware of the risks they are being asked to run. The government seems so far to have failed to still their recalcitrance.

The underlying difficulty is that there is a genuine conflict of interest between the government and the staff at the laboratory. The government, in its role as employer, most probably adequately represents the national interest in holding that applied research laboratories providing a service for customers need not be a part of the public service. But those who work in these laboratories and who have become accustomed to the apparent but often illusory benefits of being civil servants cannot be expected spontaneously to subordinate their personal interests to the greater good. The most immediate fear, at the Hydraulics Research Station but also in other laboratories where privatization threatens, is that the private enterprise will not need, or will not be able to afford, all the staff now working for them. Another problem is that public servants, unlike employees in private concerns, know in advance what their financial prospects are. The government cannot hope without trouble to spin off this or any other laboratory without compensating those directly concerned for the risks they are expected to run. There should be a generous promise of support, over a period of some years, for the core of basic research essential to commercial success. To meet the proper fear of redundancy, the government should undertake that members of staff at the laboratory will not in, say, the next seven years be worse off than if they had remained in the public service. And, to sweeten the pill, arrangements should be made that, in the immediate future, members of the staff should actually be a little better off. For if the objective is to wean public servants to life in the commercial world, is it not in everybody's interest that they should be helped to appreciate the benefits of an equitable bargain?

Supersonic scandal

Ironically, the successful landing of the United States shuttle in the Mojave Desert last week coincided with the publication in Britain of a powerful economic case against the supersonic Concorde aircraft. At this stage, the shuttle is far from being an economic proposition. The cost of development (nearly \$10,000 million) is already substantial, it is unlikely that the National Aeronautics and Space Administration will seek to recover any substantial part of this in the charges made to the ultimate users of the spacecraft and, of course, unforeseen snags may add still further to the cost of the whole development programme. But the shuttle has promise, and probably a future. Concorde, by contrast, is a technological project with a past but no future. The House of Commons Industry and Trade Committee last week published figures showing that Concorde is at present uneconomic and likely to become even more of a drain on the

financial resources of its operator, British Airways, as fuel costs increase. When the British government is skimping on public expenditure in all its works, Concorde's capacity to keep flying must be an open and urgent question.

Since its inception in 1962, the Concorde project has been an assault on economic reality. The then Mr Peter Thorneycroft (now Lord Thorneycroft and chairman of the British Conservative Party), as a member of a British government seeking entry to the European Community, seems to have persuaded himself that an Anglo-French collaboration on the development of a supersonic aircraft would have political advantages outweighing the economic costs of the enterprise. In the event, Concorde did not overcome General de Gaulle's objections to British membership, but it was left to a later British government to discover that the Anglo-French agreement about Concorde was legally indissoluble. Technically, the promise of the aircraft was always limited. Concorde was conceived of as a first modest venture into the supersonic regime for a passenger aircraft. As an economic proposition, it was too small and too slow. Optimists, two decades ago, brushed aside these faults by saying that the first Concorde would quickly be followed by other economic aircraft. In the event, there have been no successors — or even serious suggestions that there should be. The British and French governments have between them spent £1,200 million on the development of this first aircraft, and yet must continue to subsidize their nationalized airlines to keep Concorde in the air.

The cost of keeping the aircraft in the air is still substantial. Both airlines have been forgiven the cost of developing the aircraft, and British Airways has also been forgiven the £160 million cost of buying six aircraft (one of which is now out of service). Yet even though the cost of a seat on Concorde is 20 per cent more than the cost of a first-class fare on a conventional aircraft travelling the same route, British Airways is able to come within sight of breaking even only on the route between London and New York. On the whole of its Concorde operations, the airline reckons that it will have to spend £65 million more than it earns from them in the next five years. The airline told the committee that some advantages accrue to it from being seen to have a continuing interest in supersonic flight, a calculation of the intangible that does not apparently prevent British Airways asking the British government for an annual subsidy for its Concorde operations. The committee concludes that the case for Concorde is so shaky that there should be commissioned an independent financial investigation of the continued operation of the aircraft. The surprise is merely that the committee did not recommend, on the evidence that it had gathered, that the airline should turn its back on Concorde before more money has been wasted on the enterprise. Why should taxpayers, or other airline users, continue to subsidize an enterprise without financial or technological promise?

The issue is deeper than that of how a nationalized airline manages its internal affairs. To the extent that supersonic aircraft, Concorde in particular, have become identified in the public mind with the process of innovation, the continued operation of a tiny fleet of loss-making aircraft can only injure the reputation of technology as a whole. The figures assembled by the Commons committee show that the average cost of carrying a passenger from London to New York by Concorde (£543) is more than three times the corresponding cost by conventional wide-bodied aircraft. Although a substantial part of the extra cost is that of fuel — in a sense, bad luck — the poor utilization of the aircraft (on average, 2.6 hours a day) and the high cost of persuading people to buy tickets on the machine also help to make Concorde operate in the red. These extra costs will hang like an albatross around the neck of British Airways for as long as it persists in flying an uneconomic supersonic aircraft across the Atlantic. For as long, British taxpayers will be asking what the British government can intend by exhorting everybody in sight to productive industrial innovation when it chooses to spend part of its income from taxation on an unproductive innovation such as Concorde. The enterprise, wrongly conceived twenty years ago, should now be brought speedily to an end.

Telecommunications monopoly eroded?

Profiting from leased lines made seemly

The most serious challenge yet to the British telecommunications monopoly came last week in a recipe for permitting outside access to the public network. A report by Professor Michael Beesley, *Liberalization of the use of British Telecommunications Network*, commissioned by the Department of Industry, argues that private persons and companies should in future be allowed to set up separate networks but also to market telecommunications services based on circuits rented from British Telecom.

In passing, the report provides a unique account of competition between telephone monopolies and those who use their circuits for marketing private services, notably in the United States. Thus MCI Inc., one of the survivors of the companies set up in the 1960s to provide data transmission services using lines leased from AT&T (the telephone company), told Beesley that 91 per cent of its business now stems from voice and not digital transmission.

In principle, the Beesley report is intended to allow the British government to fill in some gaps in the Telecommunications Bill, still being discussed in the House of Commons. When the bill was published last October, the Department of Industry promised that provision would be made for private interests to lease lines from British Telecom, reselling the services thus made possible. In practice, opposition from British Telecom — and the sheer unfamiliarity of the Beesley proposals on the British scene — may postpone these developments.

Jargon for these services is VANS, or "value added network services". Beesley concludes that in the domestic British market there should be no restriction on the use made of British Telecom circuits by those who rent them. British Telecom, the report says, should be required to publish the costs of renting circuits of different kinds, and the charges made should be unrelated to the proposed uses. (British Telecom's proposal that it should charge royalties on VANS is rejected.) British Telecom itself should be free to compete in fields such as data transmission. Ultimately, Beesley argues, liberalization of the monopoly should be extended to the international market.

The crux of the report is Beesley's rejection of the British Telecom protest that liberalization would allow private users to "skim the cream" off its

telecommunications market. British Telecom itself has estimated that unrestricted use of its domestic circuits would lead to revenue losses of between £30 million and £110 million, and that one consequence would be slower development of the telecommunications network in Britain. Beesley doubts the arithmetic (which implies a faster growth of VANS services in Britain than in the United States), says that the public monopoly could safeguard its position by adjusting tariffs more accurately to costs and argues that the partitioning of the telecommunications market is not, in any case, a "zero-sum game". The sale of VANS would increase the amount of business to be shared by British Telecom and those selling services based on leased circuits.

As yet, Beesley finds, there is no clear pattern of how British private users would exploit the freedom to use British Telecom circuits. The most obvious starting point would be the provision of long-distance telephone services, perhaps of lesser quality, more cheaply than British Telecom. But several organizations with existing private networks — one of them the nationalized industry British Rail — have ambitions to make more use of their existing telecommunications networks.

Although the Department of Industry, the department responsible for telecommunications, favours the principles underlying Beesley, the process of consultation is expected to take several months. Political and trade union opposition is likely to be fierce.

Postgraduate medical schools in danger

The University of London's postgraduate medical schools may be the first British academic institutions to bite the dust in the wake of government cutbacks in funds for higher education. A meeting between their representatives, Lord Annan (the vice-chancellor), the Secretary of State for Education and Science and the Health Minister, before Easter, did not secure more cash but did, apparently, convince the government of a problem that it had not known existed.

At risk are the 13 members of the British Postgraduate Medical Federation and two unassociated schools, the Royal Postgraduate Medical School and the London School of Hygiene and Tropical Medicine. Their problems began last October, when government support for new overseas students was cut. The decision that such students should pay "economic" fees has since been exacerbated by reductions of central funds for universities.

The members of the British federation, with 44 per cent of their students coming from overseas, decided to charge the minimum recommended fee of £5,000 even though their average costs are considerably more. A further problem is that a fair proportion of their students are salaried staff whose fees are waived, for which the schools now receive no offsetting compensation. The result is a substantial deficit which this academic year the federation may just be able to balance with help from central funds.

The fear now is that matters will get worse. The intake of overseas students is expected to fall considerably next October. Already several governments, in particular those of Malaysia and Hong Kong which traditionally send many medically qualified students to study in Britain, have said that in future they will be sending students to cheaper universities in other countries. If the expected shortfall

materializes, some of the federation's institutes may find themselves almost bankrupt by the end of the 1981–82 academic year.

That, according to Mr D. Innes Williams of the British federation, would have "disastrous consequences" for teaching and for London's health services. Most teaching and research would have to stop and the clinical staff would be absorbed into hospitals. One possibility, that could keep the institutes functioning although crippled, would be that the government should make up the costs after fees from overseas students have been deducted.

Another possibility, is that some of the smaller institutes could become departments of the major teaching hospitals. Discussions to incorporate the Institute of Dermatology into St Thomas's Hospital are already under way. The larger institutes, however, are hoping to manage on their own.

Although there are many similarities, each institute seems to have its own peculiar problems. The Institute of Basic Medical Sciences, for example, is considering merging with the Royal College of Surgeons, which would have to put up considerable funds to bail it out of its present difficulties.

The Institute of Cancer Research, which is affiliated to the university but receives no funds from it, is suffering not so much from the shortfall in overseas students as from a decision taken in 1977 by the Medical Research Council and the Cancer Research Campaign, which provide it with the bulk of its funds, to reduce their block grant. The institute is now in the position of having spent most of its reserves to make up the deficit and of having its block grant cut even further a year from now. Plans are already under way to persuade staff to retire early. Teaching may be cut in order to retain an adequate research base.

Judy Redfearn

Environmental asbestos

Dispute on levels

Brussels

The European Commission's proposed directives on asbestos have come under attack from all sides. Many of the criticisms put forward by the European Parliament and environmental groups relate to the tougher restrictions on crocidolite (blue asbestos) than those for chrysolite (white asbestos). The dispute highlights the present uncertainty as to what can be considered a "safe" exposure to asbestos fibres.

Following the adoption of a framework directive on asbestos in 1976, the Commission has been working on four implementing directives. Two have already been put forward: one on the marketing and use of certain dangerous substances and preparations; and the other dealing with the protection of workers from occupational exposure to asbestos.

The draft directive on marketing would ban crocidolite except in the manufacture of asbestos cement pipes, acid-resisting seals, gaskets and gland packings. The marketing and use of chrysolite and other asbestos fibres would be banned for thermal and acoustic insulation, air filtering and roadway surfacing unless "the harmful use of fibres is prevented".

The directive on occupational exposure has provoked the greatest discussion. This sets limits on the concentration of asbestos fibres in the air, together with a ban on asbestos spraying, access restrictions, the keeping of health records and a commitment to introducing suitable substitutes as soon as these are available.

The Commission's proposal assumes that crocidolite is five times more dangerous than other types of asbestos. The maximum concentration allowable for crocidolite is 0.2 fibres per millilitre of air, compared with 1 fibre per millilitre for all other types.

Evidence presented at a recent international symposium on the biological effects of mineral fibres supports the Commission's position. Although earlier research with animals had shown that fibres of the same length and diameter cause the same number of mesotheliomas in the lung, more recent research suggests that human lungs react differently to different types of asbestos. Chrysolite enters the human lung less easily than crocidolite and clears more rapidly.

Some members of the European Parliament's Environment and Public Health Committee feel the evidence is inconclusive, and others feel it is confusing to have two different limit values. The Commission thinks that if a single limit value were to be adopted, it should be at the lower level of 0.2 fibres until further research is carried out. But other committee members favour using the upper level of 1.0 fibres. **Jasper Becker**

Spinning off research

Further reductions in the scale of in-house research carried out within the British Department of Defence are promised in the annual "Statement of the Defence Estimates", published last week. In a brief section on research, the white paper says that, in the years ahead, further efforts will be made to place design and development contracts with industrial companies.

This policy, recommended in a study begun in 1979 under Lord Strathcona, Minister of State at the Department of Defence, also recommended that private industry should be given (on a contract basis) more responsibility for the provision of technical and other services to the British defence research establishments. The white paper records with pride a 15 per cent reduction of research manpower between 1974 and 1980, one third of this in the last two years of the period.

At the same time, the ministry is planning to switch resources within its research establishments to long-range studies of various kinds. It cites as one promising field of research a programme now under way to develop an anti-tank weapon in which a single projectile will release a clutch of separate missiles independently guided (with the help of microprocessors) towards separate tanks.

Space shuttle

Bonding problems

Washington

As the US space shuttle achieved its first orbital test flight last week, the National Aeronautics and Space Administration (NASA) was working on proposals for an alternative thermal protection system involving large carbon/carbon panels rather than the ceramic tiles used so far.

The tiling system has been the Achilles' heel of the whole shuttle programme. Difficulty in fixing the tiles securely to the aluminium skin of the orbiter was largely responsible for the two-year delay — and corresponding cost overruns — in the first launch, and the problems are still not solved.

During last week's flight, seventeen of the protective tiles covering the engine housing on the upper surface of the shuttle became partially or totally unstuck during the launch. NASA maintained optimism, saying that none of the more important tiles on the bottom of the shuttle had been damaged, and fortunately this optimism proved justified when the shuttle returned safely to Earth.

The problems that have plagued the development of the present heat protection system — which requires 31,000 tiles, each between six and eight inches square, to cover 70 per cent of the spacecraft's surface — have recently led the space agency to



Shuttle tiles, a sticky problem

look closely at different systems which might be used for future orbiters.

During re-entry, the wing tips and bottom surfaces of the shuttle are required to withstand temperatures of up to 1,370°C. In missiles and other conventional spacecraft, such temperatures have been tolerated by the use of an ablative heat shield which burns away during the heating; but as the space shuttle depends for its cost-effectiveness on reusability, a different solution had to be found.

Using what was the most advanced technology in the early 1970s, the shielding was designed to be provided by tiles of low-density high-purity silica fibre — sometimes referred to as "foamed glass" — made rigid by ceramic bonding. There are about 20,000 of these tiles on the bottom of the Columbia, the average size being about six inches square, and the thickness varying between 0.5 and 3.5 inches.

Each tile is bonded to a pad made of Nomex felt, and the total composite skin can heat up without placing stress on the silica, which is relatively easy to repair.

The ceramic tiles have the disadvantages, however, of being both brittle and not very strong. Furthermore, the aerodynamic stresses on the surface which tend to pull the tiles from the felt which binds them to the shuttle skin had been underestimated — whereas the bonding had originally been required to withstand pressures of up to seven pounds a square inch, in practice the stresses turned out to be twice as high.

The situation became particularly embarrassing when some temporary tiles were lost during the shuttle's transportation flight from California to the launch site in Florida two years ago. And subsequent weaknesses revealed by an intensive programme of "pull-testing" required a lengthy process of densification of many thousand tiles before NASA considered that they were safe.

In the light of these difficulties, NASA last year awarded a contract to Rockwell Corporation Inc. in Downey, California, for a study of alternative thermal protection systems. Three alternative systems

are being studied, each involving a combination of three different forms of protection. For temperatures less than 540°C, insulation would be provided by panels of titanium facing covering an insulating honeycomb. In the 540°C to 980°C range, the same arrangement would be used, this time using a superalloy known as Inconel 617.

The most significant change would be in the high temperature panels, where heat shielding would be achieved by panels of reinforced carbon/carbon, up to three feet square. Various insulating materials would be placed between the shielding and the aluminium surface. The panels would be mechanically attached to the shuttle, avoiding the bonding problem associated with the ceramic tiles. Carbon/carbon is already used in limited quantities for the nose-cap and the leading edges of the shuttle's wings.

But however the new systems perform, there is still the problem of cost to be resolved. Moving to a new system would be expensive, and NASA may not be able to afford a switch.

David Dickson

Nuclear reactor safety

Go for Super-Sara

Brussels

The months of deliberation in Brussels on the future of the Super-Sara project have ended with a decision to press ahead on the lines of the Commission's original proposal. The EEC's joint research centre at Ispra in Italy will now be able to start the second phase with funds totalling 50 million European Units of Account (£0.54 = 1 EUA).

The Super-Sara project will use the experimental ESSOR reactor to study loss of coolant and other accidents in water reactors. More than 200 staff have been left in limbo since last November when the member states first began to hesitate about releasing the bulk of the project's money. In March 1980, 3.31 million EUA was granted for feasibility studies after the project was agreed in principle. But when it came to making available the 40 million EUA for the major part of the project, Germany and others began to have cold feet. The negotiations over the past six months left Germany as the only member state withholding agreement. Germany has now given way to pressure from its partners but has insisted that the project be kept below its budgetary ceiling.

However, the European Commission's initial cost projections have had to be raised since they were first made last year. Automatic salary increases and rising equipment costs mean that 64 million EUA needs to be set aside to take the project up to 1983, when it is due to be reviewed.

It seems increasingly likely that it will overshoot its schedule by as much as two years.

Jasper Becker

Election in Israel

New deal for science?

Rehovot

If, as expected, Labour Party leader Shimon Peres becomes Israel's next Prime Minister after the election in June, he will be the first man with a good grasp of science and technology to hold that office. Although not a scientist or engineer himself, Peres was responsible for initiating many large-scale research and development projects undertaken by the Ministry of Defence, where he served from 1952 to 1977, becoming minister in 1974.

These projects contributed substantially to Israel's military capabilities and also laid the groundwork for its sophisticated aircraft industry (with a turnover in 1980 of \$500 million) and for its rapidly developing electronics industry. Indeed, a very large percentage of Israeli exports based on local research (an estimated \$1,000 million last year) were spin-offs from defence research set in motion by Peres.



Peres, Prime Minister in June?

Peres has placed heavy emphasis on science-related projects in outlining the goals of a Labour-led government. There are plans for a huge "science city" in the Galilee, heavy government subsidies for energy research, a continuation of the recently inaugurated Mediterranean-Dead Sea Canal project for hydrological power and, significantly, the establishment of two nuclear power stations in the Negev.

Israel already has two research reactors, a 5-MW facility supplied by the United States and set up just south of Tel Aviv, and a more controversial 24-MW reactor from France, located near the Negev town of Dimona. But, despite an absolute dependence on imported fossil fuel for the generation of electricity and the existence of the necessary nuclear technology, Israel still has no nuclear power stations.

This anomalous situation arises primarily because the United States, the most obvious supplier of such power stations, is prevented by the Anti-Proliferation Act from selling them until

Kamikaze to Halley

The European Space Agency (ESA) is hoping to enlist the help of the Soviet and Japanese space agencies in sending its Giotto spacecraft on a kamikaze mission to Halley's comet in 1986. ESA would like to send Giotto to within 200 km of the comet's nucleus to observe molecules evaporating from it before they recombine or are interfered with by the solar wind. Initial plans had assumed a closest approach of 1,000 km because of the difficulties of pinpointing the precise position of the comet in advance and because of the risk to the spacecraft of cometary dust. If Giotto gets within 200 km of the comet it may manage a couple of hours of observation before being destroyed.

Such a close approach can only be made, however, with help from the Russian space agency which is also sending a spacecraft to the comet. The hope is that data from the Russian craft, due to arrive at the comet four days before Giotto, could provide precise details of the comet's whereabouts and allow last-minute corrections to Giotto's course. The Russian spacecraft itself will not be able to approach nearer than a few thousand kilometres to the comet nucleus because, unlike Giotto, it lacks a dust shield.

The Japanese space agency will also be sending a spacecraft — at 135 kg the smallest of the three — to photograph the comet in the hydrogen Lyman alpha line. The three space agencies are discussing how they might exchange data gleaned from their separate missions, and next summer the three, together with the US National Aeronautics and Space Administration (NASA), will be meeting to discuss a collaborative project between all four agencies. Although NASA has no mission of its own to Halley's comet, it is organizing a "Halleywatch" which will make use of ground as well as space based observations.

Judy Redfearn

Israel signs the nuclear non-proliferation agreement and opens its nuclear facilities to inspection by the International Atomic Energy Agency.

It was only after Egypt signed this agreement in February that the United States agreed to provide that country with two nuclear power stations with a combined capacity of 2,000 MW. But Peres refuses to say whether a Labour-led government will do what all previous governments have refused to do, namely sign on the dotted line, or, if not, how he expects to get the reactors he proposes to place in the Negev.

Peres sees a close link between the country's economic development on the one hand and its scientific and technological development on the other. He accuses the present Begin government

of allowing the economy to stagnate and, as a result, setting in motion a brain drain that has deprived Israel of thousands of scientists and engineers.

But many specific questions remain unanswered. Will, for instance, Labour create a Science Authority or even a Ministry of Science to take over from the virtually moribund National Council for Research and Development?

Also unclear is the amount of government support that can be expected for basic research, hitherto dependent on overseas grants, and whether there will be more funds for universities, which have suffered a drop of more than 30 per cent in real income over the past few years.

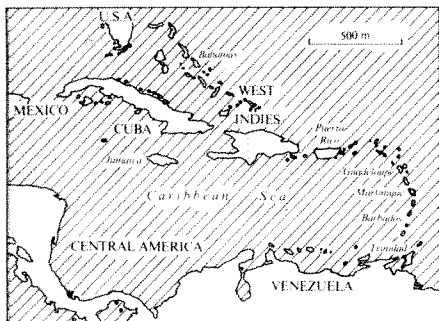
Peres will say no more than that he plans to "upgrade science, technology and education". How this will be reflected in Labour government budgets remains to be seen.

Nechemia Meyers

Caribbean anti-pollution plan Jamaican entente

Montego Bay, Jamaica

Temporarily forgetting their many political and cultural differences, government officials from 23 Caribbean nations met in Jamaica recently and agreed on guidelines for coordinating efforts to minimize the pollution of their seas and coastlines. The meeting was the latest in a series organized under the Regional Seas Programme of the United Nations Environment Programme (UNEP), most widely publicized of which have been those aimed at cleaning up the Mediterranean.



The plan of action agreed in Jamaica will probably have less immediate impact than the Mediterranean agreement, in part because the Caribbean is not a completely closed sea, and the level of pollution is therefore considerably lower.

Nevertheless, delegates to the Jamaican meeting stressed that the prime goal of the Caribbean plan should be preventative. According to Dr Stepan Keckes, director of the UNEP programme, "the governments have decided to tackle marine pollution not as a problem in itself but as the consequence of the problems which lie on the land".

The principal goal is to integrate concerns for the state of the marine environment into development plans for

the Caribbean region. The four projects already selected as being of top priority under the action programme include a regional oil-spill contingency plan, guidelines for managing watersheds, efforts to improve national environmental health services and broad measures to improve environmental education.

Relatively little money has been promised so far. A trust fund which is being set up as part of the action plan failed to reach its target of \$1.5 million, as a result of which UNEP secretary general Mostafa Tolba said that his organization would contribute an extra \$1.38 million. The principal donors to the fund have been Mexico, Venezuela and France, the first two being the main economic forces in the Caribbean region. Britain is prepared to make a contribution that reflects the specific interests of its dependent territories in the region (which include Antigua and the Caymans) but is unlikely to match the \$375,000 promised by France.

UNEP officials, however, as well as those from other aid bodies, hope that the plan will act as a stimulus for action at a national level. Jamaica in particular is hoping to take a leading role in coordinating development policies in the region. Addressing the closing session of the meeting, Jamaica's recently elected prime minister, Mr Edward Seaga, said that the agreement on marine pollution was "the greatest collaborative effort in the Caribbean to date, and may be a vanguard of future efforts to bond Caribbean nations in a common cause".

Jamaica is to become the base for a regional coordinating unit that will be established to facilitate the technical implementation of the action plan. The island has also put in a bid to become the home of the international sea-bed authority which would result from a successful outcome of the Law of the Sea negotiations now taking place at the United Nations in New York.

In the long term, the success of the Caribbean action plan is likely to depend primarily on procedures that individual states can be persuaded to adopt to minimize the sources of marine pollution. A top priority is to draw up a draft regional convention, and UNEP officials hope that a treaty committing countries to reducing sources of potentially damaging pollution can be agreed within two years.

The fate of the plan is also likely to be important to UNEP itself. Moves in Washington to reduce, or even eliminate, the US contribution to UNEP could cripple many of its programmes. US officials, however, are keen to cement political ties in the Caribbean — particularly with the new Jamaican government — and have indicated that the United States might in future be prepared to contribute to the trust fund. If UNEP can prove its value to the United States in the region, its medium-term prospects could be significantly improved.

David Dickson

Philosophy wrangle

Doubts persist about academic freedom in post-Tito Yugoslavia. Last month, the Yugoslav Secretariat for Information announced the reinstatement of the seven Belgrade professors of philosophy dismissed last January for their association during the early 1970s with the independent Marxist journal *Praxis*. Apparently the seven are to receive new appointments in an "Institute for Social Studies" specially created for them. Members of the group, however, say that no definite settlement has been reached.

Since *Praxis* was closed down in 1975, and the members of the group were suspended on 60 per cent salary, forbidden to teach or to publish within Yugoslavia, the authorities have repeatedly tried to persuade them to find new posts.

In June last year, the Serbian republic's law on universities was amended to permit dismissal after six months' suspension on the grounds of "damaging social interests". During the next six months, the authorities tried to persuade the suspended professors to find jobs. Two were even offered non-teaching posts within the philosophical faculty of Belgrade University, but they replied that they could not accept the post unless the other five were also reinstated.

Dismissal became effective in January, but the authorities later suggested that posts might be found in the Belgrade "Institute of Social Studies" — not part of the university but the authorities suggested that it could be rapidly converted into one.

Official plans, it turned out, envisaged that three of the group would retire, and the other four would be scattered among the existing centres of the Institute of Social Studies, covering areas such as "public opinion" and law. When these plans became known to the seven, they refused emphatically: they were, they said philosophers, and wished to work as such.

The vice-premier of Serbia, Milan Dragovic, who was acting as official negotiator, had already said that the transfer of the Institute for Social Studies to the university was nothing to do with the seven's demands, but was something that had long been necessary. He also conceded that a Centre of Philosophy within the institute would be a good idea.

On 3 April, the acting director of the institute, Dr Svetozar Culibrk, called in the whole group for talks, the first time that the authorities have been willing to negotiate with the professors *en bloc*. Dr Culibrk said that he "liked the idea" of a centre and would lay it before the institute's council. The seven are understood to be cautiously optimistic.

Vera Rich

CORRESPONDENCE

Irish evidence

SIR — Mr Peter H. Roberts's letter (*Nature* 19 March, p.184) offers the importation of infected Irish cattle as an explanation for the fluctuating incidence of bovine tuberculosis in the South West region between 1974 and 1980. It also says that the Zuckerman report carefully avoids this factor. I should like to set out the facts.

Table 10, on page 57 of the report, records the ministry's assessment of the sources of bovine tuberculosis infection in all infected cattle herds in Great Britain during 1972–1978, including imported Irish cattle. The figures for the South West and Sussex show 331 cases attributed to badgers and 4 (of which 2 were in Sussex) to imported Irish cattle. The totals for the rest of Great Britain are 0 and 171 respectively. It is little wonder therefore that Lord Zuckerman did not pursue the Irish cattle question when investigating infection in badgers in the South West.

The following breakdown of the figures for infection attributed to Irish cattle imports further detracts from the force of Mr Roberts's argument

	Great Britain	SW England
1972	59	1
1973	23	0
1974	22	0
1975	23	0
1976	23	1
1977	12	0
1978	13	0
TOTAL	175	2

Traditionally the Irish trade has been, and still is, mainly with the North-East of England and East Scotland. In the 2-year period before 1976 when, Mr Roberts suggests, cattle infected with bovine tuberculosis were imported from the Irish Republic, there was no increase in disease incidence from this source. What the figures do show is the effectiveness of the decision to introduce, in 1976, pre-export testing of Irish cattle.

W.H.G. REES

Ministry of Agriculture, Fisheries and Food,
Tolworth, Surrey, UK

Who is Nabi?

SIR — Readers may wish to know that the name of Isadore Nabi, the signatory of a recent letter criticizing my views on sociobiology and ethics (*Nature* 19 March, p.183) is fictitious. Should the writer ever make a statement over his own name, I hope he will confess that he lifted the two 1975 phrases of mine out of context in a way that reverses the meaning of one and makes it appear to contradict the other. I also trust that he will mention my later and fuller treatments of sociobiology and ethics in *On Human Nature* (1978) and *The Tanner Lectures on Human Values, Volume I* (1980).

EDWARD O. WILSON

Museum of Comparative Zoology,
Harvard University, Massachusetts, USA
Isadore Nabi is believed to be the pseudonym of Professor R.C. Lewontin of Harvard University — Editor, *Nature*.

Erratum

An incorrect spelling was given for the oak wilt fungus in the 26 March issue (*Nature* 290, p.284). The organism is *Ceratocystis fagacearum* and the beetle involved in its transfer is *Scolytus*.

Darwin's truths

SIR — The letter from your correspondents at the British Museum (Natural History) (*Nature* 12 March, p.82) relating to your editorial "Darwin's death at South Kensington" (26 February, p.735) fully supports your contention that something is amiss at that institution. The writers do not appear to understand the difference between a theory and a fact. When Darwin's *Origin of Species* was published in 1859, it was presented as a theory. It was then, and for long afterwards, proper to refer to it as "the theory of evolution". But since then the evidence for the theory has accumulated from many different sources and in many ways, among them natural and laboratory experiments. Wherever and by whatever means the theory has been tested, it has withstood every attempt at falsification.

The atom at the beginning of this century was a theory. No one today will doubt that it is a fact. Yet no one has ever seen an atom. Yet we have seen evolution in process before our very eyes in heritable changes in many forms, perhaps the most remarkable and obvious is the well-known case of industrial melanism in moths. What kind of proof do the South Kensington writers require before they will be willing to accept evolution as a fact? There are many differing theories concerning the mechanisms of evolution, and these are all to the good, but the fact of evolution as a process of change, surely, cannot be denied. The proofs for it are overwhelmingly clear.

ASHLEY MONTAGU

Department of Anthropology,
Princeton University,
New Jersey, USA

SIR — I am at a loss to understand what all the fuss is about concerning the phrase "if the theory of evolution is true". Darwin used it. I quote from *Origin of Species*:

"There is another and allied difficulty which is much more serious. I allude to the manner in which species belonging to several of the main divisions of the animal kingdom suddenly appear in the lowest fossiliferous rocks . . . If the theory of evolution be true it is indisputable that before the lowest Cambrian strata was deposited long periods elapsed as long or probably far longer than the whole interval from the Cambrian to the present day, and that during these periods the world swarmed with living creatures . . . The difficulty of assigning any good reason for the absence of vast piles of strata rich in fossils beneath the Cambrian system is very great."

This, coupled with the following quote by Dr W.R. Thompson, Fellow of the Royal Society, in the foreword of the *Origin of Species* (1956) does nothing to inspire confidence or belief in the current theories of how life started or progressed on Earth.

"It does appear to me in the first place that Darwin in the *Origin of Species* was not able to produce palaeontological evidence sufficient to prove his views, but that the evidence he did produce was adverse to them, and I may note that the position today is not notably different . . . As we know there is a great divergence of opinion among biologists not only about

the causes of evolution but even about the actual process, the divergence exists because the evidence is unsatisfactory and does not permit any certain conclusions."

I am left wondering which theory requires the most blind faith.

GORDON SMITH

Liverpool, UK

Room for all

SIR — At first amused, I am now saddened by the "Death of Darwin" controversy (*Nature*, 26 February, p.735 *et seq.*). We have slipped back a hundred years: how long before letters signed Wilberforce and Huxley appear? Why "either-or"? Why cut off either right or left hands? Is there not room for Darwin, Hennig, God — and even Marx?

J.R. BAKER

Cambridge, UK

Popper's philosophy

SIR — Accusations that museum displays organized in the light of cladistic philosophy represent creeping Marxist-Leninism were bad enough. But when *Nature* sees fit to defend the scientific status of Darwinism on the grounds that "metaphysical theories are not necessarily bad theories"¹ then matters philosophical have truly gotten out of hand.

It should come as no surprise to those familiar with the writings of various cladists that the writings of Professor Popper on evolutionary theory are held in great esteem. After all, he has argued for years that no view of history, whether human, biological or technological, admits of the classification scientific. For Popper the events of history are unique and, thus, not amenable to systematic explanation via theories of any sort². Since cladists find the infusion of the least amount of theoretical insight into classification suspect in much the same way that their positivistically minded cousins, the phoneticists, did a decade or so ago, it is hardly surprising that they would greet the rediscovery of Professor Popper's writings from the 1940s and 1950s with great glee.

Popper's hostility to evolutionary theory explains the favour his views receive from systematists fond of cladism. But it does not explain why this same hostility should inform the judgments others make of the scientific status of Darwinism. Popper's views have been roundly and soundly criticized by numerous philosophers interested in the scientific status of evolutionary theory. The apparently timeless contentions that the theory is (1) tautologous, (2) unfalsifiable, (3) lacks predictive power and (4) lacks confirmation have shown to be false by a decade of scholarship in the philosophy of biology dating from the appearance of David Hull's *Philosophy of Biological Science*^{3,4}. Why *Nature* should choose the cladists' favourite philosophical authority over the ruminations of contemporary authors gives one pause not about politics but about the reading habits of the scientific community.

There are numerous reasons for not taking Professor Popper's criterion of demarcation of science and non-science seriously. Perhaps the most obvious is that it just does not cut well — cosmology and evolutionary theory wind up in the metaphysical hopper, astrology

and phrenology make the grade as science — false science but nevertheless science.

Numerous other contributors to the philosophy of science such as Nagel, Hempel, Hanson, Toulmin, Kuhn, Shapere, Lakatos, McMullin and many others have shown serious conceptual flaws in Popper's defence of the falsifiability criterion for demarcating science from non-science⁵.

Since the issue is so badly mauled in the *Nature* editorial let us take up the matter of the falsifiability of Darwinism since it seems of such great concern to cladists, creationists, and correspondents to this journal. The issue is stated as the question, "Is Darwinism falsifiable?" admits of a simple answer — yes.

If we are talking about the version of evolutionary theory propounded by Darwin in the *Origin*, then that theory has been shown false many times over. It claimed that all organic variation could be accounted for by natural selection and a tendency to inherit among all creatures. This is false. Mutation and recombination at the level of genes refute the adequacy of Darwin's posited mechanisms. Darwin claimed that the bulk of speciation occurred through anagenesis — evolution within a lineage. This is false. Cladogenesis or the splitting of an interbreeding population into two groups via natural barriers or other means accounts for the bulk of speciation. Darwin claimed that social behaviour exists in insects as a consequence of group benefit. This is false. Social behaviour exists among many insects as a consequence of kin selection, reciprocal altruism, parental manipulation or some combination of these mechanisms. This list could go on. The point is that evolutionary theory, like all theories in science is constantly tested, refined, modified, adapted and re-written. The issue confronting defenders of the validity and utility of evolutionary theory in its early or modern forms is not whether it is science or not — Darwin himself certainly settled that matter with admirable skill in the empirical and analogical evidence he brought forward in his own writings. Rather the question is how does one go about invalidating scientific theories that constantly change and evolve — an issue that ought to attract the attention of the cladists if not the creationists.

How true is the theory of evolution? It is as true a theory as there is in science. Which is to say it depends on who is asking and when. Of course the theory of evolution is not a fact. Evolution is a fact and the theory of evolution tries to explain this fact. But the theory of evolution surely passes muster as science, even on the out-dated grounds cited by Professor Popper in his early work. (Even he no longer believes in the adequacy of his early analyses of evolutionary theory⁷.) The real question that should concern scientists is whether they know enough about current thinking in the history and philosophy of science to know a sound theory when it stares them in the face — a circumstance unlikely to occur at the British Museum unless this ignorance is alleviated.

ARTHUR L. CAPLAN

The Hastings Center,
New York, USA

1. *Nature* 290, 75-76 (1981)
2. Popper, K. R. *The Poverty of Historicism* (Routledge & Kegan Paul, London, 1957).
3. David Hull, *The Philosophy of Biological Science*, (Prentice Hall, Englewood Cliffs, 1974).
4. Ruse, M. *Philosophy of Biology* (Hutchinson, London, 1973).
5. Suppe F. (ed.) *The Structure of Scientific Theories* 2nd edn (University of Illinois Press, Urbana, 1977).
6. Caplan, A. L. *Erkenntnis* 13, 261-278 (1978).
7. Ruse, M. *Philosophy Sci.* 44, 638-661 (1977).

Academia stagnant?

SIR — I see no reason for your concern (*Nature* 19 March, p.175) about the ability of British universities to recruit young scientists to academic posts. The present policy of recruiting young persons in their mid-twenties and sometimes without postdoctoral or international experience has so little merit that nothing would be lost if this recruitment of lecturers was totally ended. The best of our young scientists will wish to stay longer in full-time research and will continue to be available to the universities for years to come.

The rapid recruitment of academic staff in the 1960s expansion phase is unlikely to be the cause of stagnation in science departments. Where stagnation exists it is fairly certain that the department has defective leadership which might very well result from easy promotions gained in the years of expansion. Universities should, especially when resources are tight, recruit scientists into permanent teaching posts later in their careers than has been customary in Britain. A changed policy could provide continuing opportunities for the present generation of young scientists. Retirements and early retirements will improve some departments and at the same time allow new entrants or well justified internal promotions which ought to maintain the intellectual vitality of any department worth keeping.

J.R. PENSWICK

School of Biological Sciences,
University of East Anglia, Norwich, UK

Freedom from NATO

SIR — With the mounting concern in your columns on measures which scientists should take to ensure academic freedom for colleagues deprived of their rights, it is surprising that you publicize the NATO Advanced Study Institutes, that scientists are happy to perform on the stage which NATO sets, and that no mention of this paradox has been made in your correspondence.

No doubt these conferences fulfil a useful scientific function in promoting exchange of information in very congenial surroundings. Under normal circumstances, the argument that funds (from a rather dubious source) were being channelled to a good cause would have been sufficient to rationalize the situation. In some cases there is also the hope that the opinions of the scientific community will filter back along this channel, and that contact with sensitive institutions will lead to a more extensive dialogue and to a greater opportunity for diplomatic manoeuvres. Often, such occasions bring us into contact with scientists who have difficulty in meeting foreign workers because of governmental restrictions.

These conditions do not apply to attendance at NATO-supported conferences. There is little evidence that the meetings provide a forum for discussion with NATO officials in any way which would allow scientists to exercise their social responsibility. Nor could it be claimed that the activities of the conferences could not easily take place under other auspices. But most seriously, at a time when nuclear installations in Europe, products of scientific ingenuity, threaten the peace of our own and other countries, this close relationship between scientists and NATO can only be seen from outside as complete approval for (if not connivance with) the increasingly military orientation of our society.

This association with NATO must also affect the scientists themselves. After participation in any symposium it is difficult to leave without feeling some debt of gratitude to the organizers who lavish money and attention on their charges. Qualms about the ethics of the institution running it are henceforth stifled. Thus the NATO scientific conference is a highly effective public relations exercise both for the participants, and those who look to scientists for a lead in problems created by technology. How can Soviet colleagues be expected to take our views seriously when they see the strength of our opinions so cunningly controlled by political agencies? If our protests are to carry conviction in any field, we should dissociate ourselves from the influence of organizations like NATO, and take every opportunity to make this known publicly. ROGER R.C. NEW
University of Liverpool, UK

Conservation sites

SIR — Many criticisms can be levelled at the system of Sites of Special Scientific Interest (SSSIs) but Muir's arguments (*Nature* 12 March p.82) are untenable.

Extinction may be "normal" but that is irrelevant to "value". To analogize, we will all die, but do not usually accept this as an excuse for murder. We are replacing complex ecosystems with much simpler landscapes, allowing for fewer species to exist or appear.

In a limited sense, rarity may support a value judgement of "biological deficiency", and the mere presence of a rarity may be poor grounds for conservation; but rare species can be good indicators of environmental quality when considered along with other evidence.

SSSI status may result in a "loss of freedom" of the owner. Losing such sites is a far greater "loss of freedom": it destroys our options and those of future generations, for we cannot replace them.

It is also misleading to allege just that protection of SSSIs is a cost to their owners: it prevents owners *increasing* their income by reclaiming the site, but that increase would itself be a cost to the taxpayer in terms of grants for reclamation, and later subsidies.

It is astonishing that someone from a university zoology department could contend that "Once . . . information is recorded and published . . . nothing new will be learned by preservation". No complete record of any site or habitat exists, and the same material can generate an infinite variety of "descriptions". Frequently we cannot turn to the literature for basic data to test new theories — we have to test them through fresh fieldwork designed specifically for that purpose. If we lose the field sites, ecology becomes metaphysics and not science, and to rely just on the environments we are creating today leaves us a severely restricted data base.

This necessity has fortunately been recognized in other fields. Once, excavators of bone caves destroyed vast storehouses of information by removing every vestige of material but failing to distinguish stratified deposits. Later workers recognized strata but again lost vital information by failing to extract pollen, rodent bones, and other small items. Now sections of deposit are left undisturbed for the future application of new ideas and techniques. Ecologists need similar opportunities.

KEVIN A. ROBERTS

Hoddesdon, Herts., UK

NEWS AND VIEWS

Genes of lymphocytes I: Diverse means to antibody diversity

from Miranda Robertson

AFTER fifteen years of controversy and five years of DNA cloning, immunologists are at last unanimous on how antibody diversity is generated. This remarkable eventuality is due to the accumulation over the past five years of evidence for almost every mechanism anyone has ever suggested; and to which has now been added yet another, of such baroque complexity that it could never have occurred to anyone to suggest it but for the evidence (albeit preliminary) in the DNA.

There used to be, broadly, three schools of thought, one of which adopted the extreme position that the entire antibody repertoire was encoded in a vast array of germ-line genes, and the other two of which favoured the somatic generation of variation on the basis of a limited number of germ-line genes. Two means were proposed for the somatic generation of variation. One was simple point mutation; but this on its own would not explain the distribution of the variation, which is concentrated in three "hypervariable" regions on the variable region gene (see Fig. 1). The other was recombination, represented in its most elegant and extreme form by the mini-gene theory of Kabat, who proposed that a range of genes coding for each of the three hyper-variable regions and the four less variable (framework) regions existed separately in the genome, to be brought together in variable combinations during the differentiation of the B lymphocyte. This theory however did not explain the evidence for point mutations in the framework regions. All these views have now been at least partly vindicated, as became clear at the massive immunology symposium in Paris last summer¹; and at a more recent and smaller symposium in Salt Lake City*, their proponents were still reeling with surprise at finding they no longer had anything to quarrel about.

Two things became clear very soon after the cloning and sequencing of the first immunoglobulin genes. First, there is quite a large number of germ-line genes: more than one for each of the known variable region subgroups (a subgroup is a family of

variable regions defined by similarity in the framework regions). Estimates have been made for mouse λ and κ light chain genes. Whereas the mouse λ V gene family contains only two genes², S. Cory (Walter and Eliza Hall Institute, Melbourne) now puts the size of the κ gene family at between 100 and 300; however, less is known about the possible numbers of heavy-chain variable region genes which she and her colleagues are in the process of investigating.

Second, the differentiation of a B lymphocyte is accompanied by the translocation of V region DNA. In the case of the light chain, one of the germ-line V genes, comprising the first three framework segments with two hypervariable regions and part of the third, is brought into proximity with one of a cluster of J genes comprising the rest of the third hypervariable region and the fourth framework segment. Since V genes may recombine with J genes at random, and with variation in the site of the join, this immediately generates considerable diversity. The scope for diversification is even greater in the case of the heavy-chain

genes, in which the V gene, which extends only to the end of the third framework region, recombines with a D gene comprising the third hypervariable region, and the D in turn recombines with one of a cluster of J genes coding for the fourth framework region. The D and J heavy-chain genes thus behave very much like the mini-genes postulated by Kabat; but Kabat never dreamed of the extraordinary properties of D genes that are beginning to emerge from the latest investigations of S. Tonegawa and his colleagues (Basel Institute of Immunology).

They have used a D region probe² to locate a cluster of eight homologous D genes in germ-line DNA. All eight genes are 17 base-pairs long; yet expressed D regions with substantial "core" homology to the germ-line genes range from thirteen to more than 40 base-pairs in length. This is too extreme to be explained by variations in the site of V-D-J recombination. Tonegawa has suggested that the explanation may lie in D-D joining, by a mechanism that would not result in simple tandem repetition of the homologous sequences. The mechanism depends on the

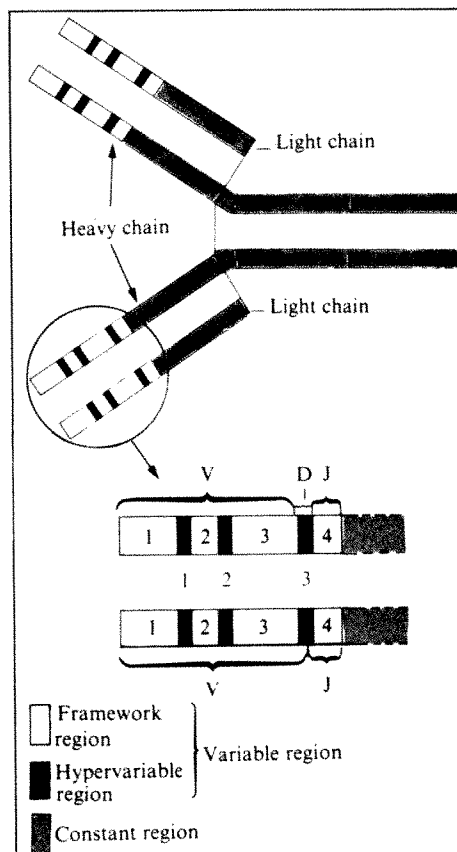


Fig 1. Immunoglobulin molecules are made up of four chains: one pair of heavy chains and one pair of light chains. Each chain is divided functionally into two regions: the variable region, which mediates antigen recognition, and the constant region, which in the case of the heavy chains mediates the effector function of the molecule.

The variable region can be further subdivided into four regions which vary only slightly from one antibody molecule to another, and are known as the framework regions, and three regions (the hypervariable regions) which are very much more variable and in the folded protein form the actual antigen-binding site.

Each light chain is now known to be coded by three separate genes: a C gene coding for the constant region, a J gene coding for the fourth framework region and part of the third hypervariable region, and a V gene coding for the first three framework regions including hypervariable regions one and two and part of hypervariable region three.

Each heavy chain is coded by four separate genes: a C gene coding for the constant region, a J gene coding for the fourth framework region, a D gene coding for the third hypervariable region, and a V gene coding for the first three framework regions including hypervariable regions one and two.

Miranda Robertson is associate editor of Nature.

*The ICN/UCLA Symposium on *Immunoglobulin Idiotypes and their Expression* was held in Salt Lake City, Utah, on February 8-15, 1981.

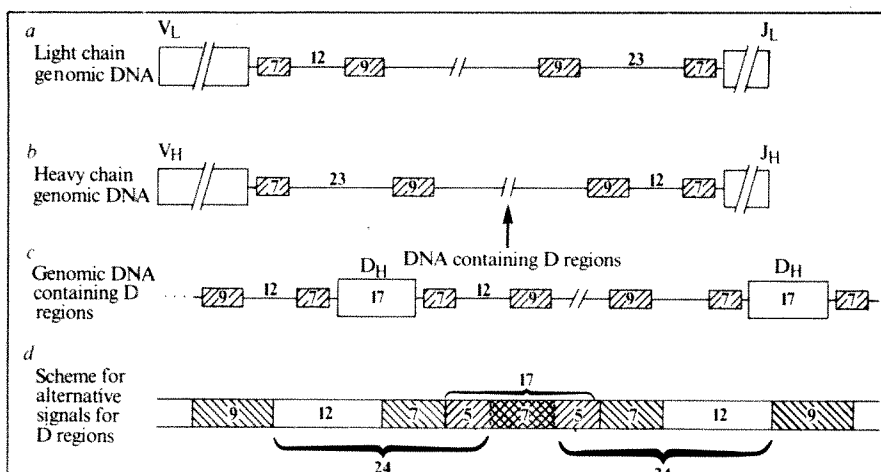


Fig. 2 The pattern of genomic sequences believed to mediate V-J and V-D-J joining in light (a) and heavy (b and c) chain variable region genes. Open rectangles are coding sequences, shaded rectangles are nonamer and heptamer signal sequences, and intervening lines are non-coding DNA. The figures 12 and 23 on the non-coding DNA refer to the length of the spacer regions between the heptamer and the nonamer, which are highly conserved sequences believed to act as recombination signals: the heptamer is followed by the nonamer after a space of 12 (strictly, 11 or 12) base pairs downstream of the V_L gene, and inversely, the nonamer is followed by the heptamer after a space of 23 (strictly, 22 or 24) base pairs upstream of the J_L gene (Fig. 2 a). Heavy-chain V genes are followed by the same signal sequences with the 23 spacer, and heavy-chain J genes are preceded by the inverse signal, again with a 23 spacer (Fig. 2 b). Early *et al.*⁴ and Sakano *et al.*⁵ surmised that recombination depends on the pairing of the signal sequences plus a 12-base-pair spacer with the inverse signal sequences plus a 23-base-pair spacer; they accordingly predicted that D_H genes should be flanked by signals with 12-base-pair spacers (Fig. 2 c). This has proved to be the case². If the 12/23 rule holds, however, it should make D-D joining impossible; and that is where the baroque complexity comes in.

Tonegawa and Kurosawa have proposed that each D region can treat the central 7 of its 17 base-pairs as the heptamer of the signal sequences and add the five coding base-pairs plus their adjoining heptamer, either upstream of downstream, to the adjoining spacer to make a 24 instead of a 12 base-pair spacer (Fig. 2 d). This would make it possible to add about 5 base-pairs of one D gene to another without violating the 12/23 (or 12/24) rule or producing tandem repeats, which are not found.

12/23 rule first enunciated by Early *et al.*³ for V-J recombination, and is described in detail in the legend to Fig. 2.

Even without D-D joining, the evidence for which is at best circumstantial, the recombination of a substantial repertoire of V genes at random with J and D genes would be sufficient to account for an antigen-binding repertoire well into the thousands. To this variation within chains must be added the variation generated by the association of different light chains with a given heavy chain. Since the expression of the heavy-chain genes precedes that of the light-chain genes, it has been assumed that each different heavy chain associates with a different light chain in the progeny of a single pre-B cell⁶. By fusing a pre-B cell line with a myeloma line that does not express L chains, W. M. Kuehl and his colleagues (Virginia University) have succeeded in inducing the expression of light chain genes of the pre-B cell and demonstrated that different rearrangements of the κ genes do in fact occur in the progeny⁷.

Finally, there is growing evidence for further diversification of the rearranged immunoglobulin genes by somatic mutation in the course of B cell ontogeny. In the absence of a B cell line whose ontogeny can be followed *in vitro*, the evidence is necessarily indirect. But the availability of cloned germ-line genes for

the V-regions of some antibodies has made it possible to compare the germ-line sequences with those of secreted immunoglobulins from B cells at different stages of maturation. At early stages of maturation, B cells synthesize immunoglobulin M; at later stages they synthesize immunoglobulins of other classes: IgG, IgA or IgE. Thus by comparing IgM and (say) IgG in turn with the homologous germ-line sequence, it is possible to detect any increase in departures from the germ-line gene that may have occurred as the cells switched from one class of immunoglobulin to the next.

Such increases have been reported both by A. Bothwell (Massachusetts Institute for Technology) and from the laboratory of L. Hood (California Institute of Technology), working with antibodies raised against nitrophenol (NP) and phosphorylcholine (PC) respectively. Antibodies raised against these antigens possess, in certain strains of mouse, the interesting and useful property that a very high proportion of the total immunoglobulin bears a characteristic antigenic determinant, or idio type, of its own (NP^b in the case of anti-NP antibodies, and T15 in the case of anti-PC antibodies). But not all even of those antibodies bearing the same idio type are by any means identical, and in the light of recent analyses this is hardly surprising. Both Bothwell⁷ and

Hood report that there is a cluster of several germ-line V genes all with substantial homology to the idio type-bearing anti-NP and anti-PC proteins. The existence of such homologous V-gene clusters (as well as Tonegawa's homologous D gene clusters) strongly suggests duplication of genes during evolution — particularly since two of the five homologous genes sequenced by Hood and his collaborators are actually identical.

As it turns out, comparison of the partial amino-acid sequences of several T15 proteins with the germ-line sequences shows that all the proteins are derived from the two identical germ-line genes, but with numerous substitutions, notably in the proteins of ontogenetically later classes (IgG and IgA)⁸. None of these substitutions can be explained except by mutation, with one exception which could (but need not) be due to recombination between two of the germ-line V genes.

The sum of all these recent data leaves us with an embarrassing riches of diversity-generating mechanisms, but some important outstanding questions about their deployment. The first and most obvious is exactly how many germ-line genes there are. It is believed that there are four functional J genes each for the light and heavy chains; but estimates of the numbers of V genes still range between 500 and 1,000, and little is known of the total size of the D-gene family, of which that sequenced by Tonegawa and his associates may represent only a branch.

The second and more challenging is whether a special mechanism, still undiscovered, is needed to explain the concentration of amino-acid substitutions in the hypervariable regions. Both germ-line sequence and somatic recombination go some way towards explaining this pattern of variation. Related germ-line genes differ from one another more in the hypervariable regions than in framework-coding regions; and variable V-J and V-D-J joining account for further variation in the third hypervariable region — particularly in the case of V-D-J joining. But Gearhart *et al.*⁸ have shown that somatic mutations too are clustered, at hypervariable regions 1 and 2.

This could in principle be explained without recourse to further special mechanisms, simply by somatic selection. B cells whose surface immunoglobulin does not bind antigen produce no progeny: thus those expressing mutations in the hypervariable regions which form the antigen-binding site will tend to be selected. Selection will also occasionally operate on mutations affecting framework sequences: although it is the hypervariable regions that actually make contact with antigen, the framework regions may influence the manner in which they do so — for example, framework variants of an immunoglobulin which in its original form bound PC both on its own and complexed to a carrier will no longer bind the PC-

carrier complex (M D Scharff, Albert Einstein College of Medicine) Both the clustered substitutions in the hypervariable regions, and the occasional ones in the framework can therefore be accounted for by selection for antigen binding

It would be a mistake however to assume that all mutations in expressed immunoglobulin genes must have been specifically selected indeed, some are silent and could never have been subjected to selection These are presumably the "neutral mutations" of the immune system But if the pattern of mutations requires no special explanation, it is possible that the rate of mutation may Certainly many immunologists are wondering about possible mutagenic reactions associated with the class switch Notwithstanding the considerable difficulties associated with the elucidation of molecular mechanisms of immunoglobulin gene rearrangement,

the issue of mutation rate may be resolved more quickly by research at this level than by arguments about sequences, because arguments about the mutation and selection of sequences depend on a detailed understanding of the relationship between the sequence of an immunoglobulin and its antigenbinding properties and this is usually unknown

However, each antibody molecule can now be seen as the end product of a series of selective steps, beginning with the germ-line genes The germ-line V, J and D genes have presumably been selected in the course of evolution for their ability to produce proteins that will bind to common environmental antigens There follow at least two rounds of somatic selection by antigen The first acts on the V-J and V-D-J recombinations that produce IgM It is followed by at least one more round of somatic selection favouring mutations that produce efficient binding to the same antigen, or indeed to any other antigen present in large enough quantities to encounter the mutant cell A detailed analysis of hybridoma proteins from immunized and hyperimmunized mice, such as Bothwell and his colleagues have already begun, may help to show how far antigenic selection could account for the range of antibodies a single mouse is able to produce against a single antigen

coupled with the missing oceanic crust, arising from Earth expansion during the late Proterozoic Pure Earth expansion has the effect of leaving continental centres in the same coordinate position and so has no effect on the polar wandering paths The problem is to see how the relative continental motions during the Phanerozoic, which palaeomagnetism brings out so clearly, could have returned the continents to the relative positions they had in the Proterozoic

The main counter-arguments to expansion were the well known length of day calculation (Runcorn *Nature* 204, 823, 1964) based on daily and annual growth increments in corals, the lack of secular change in global sea level over the Phanerozoic (Vail *et al Mem Am Ass petrol Geol* 26, 83, 1977) coupled with near stability in the volume of the hydrosphere (Anderson *Rev Geophys Space Phys* 13, 37, 1975), and my demonstration that for the past 5 years, half or more of the subduction required by plate tectonics on an Earth of fixed radius is actually occurring These arguments give a maximum rate of change in Earth radius of 0.1 cm yr⁻¹ since the Devonian, in gross conflict with the idea of accelerating expansion during the late Phanerozoic This conclusion is backed up by the provinciality of Palaeozoic faunas, pointed out by Clive Burrett, which suggests that major ocean basins existed at that time

All this, however, fails to provide any hard evidence about palaeoradius in the Proterozoic, to which attention is now focused Only Stewart's (*J geol Soc Lond* 133, 281, 1977) palaeogravity method, which compares computed geotherms with mineralogical evidence for the actual *P-T* conditions in the crust and upper mantle, gives any help here Palaeoradius can easily be calculated from the surface gravity field provided that Earth mass and the universal constant of gravitation are known Gravity seems to have remained within a factor of two of its present value throughout Earth history, implying a rate of change in radius < 0.1 cm yr⁻¹ and hence limiting late Proterozoic expansion to < 1,000 km This is too small by a factor of 2.5 to solve the problem of the missing oceanic crust

It is to be hoped that the basically sound palaeomagnetic technique for measuring palaeoradius will soon be applied to early Proterozoic basic dykes across a stable continent such as Canada Using samples collected specifically for the purpose, it should be possible to reach a definite decision about the likelihood of late Proterozoic Earth expansion Published palaeoradius estimates using the palaeomagnetic technique have always used samples from Phanerozoic rocks and even these, it was pointed out at the meeting, did not always come from stable continental areas as required by theory, so that the validity of the palaeoradii was open to doubt □

- 1 see Robertson, M. *Nature* 287, 390 (1980)
- 2 Brack, C., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* 15, 1 (1978)
- 3 Kurosawa, Y., von Boehm, H., Haas, W., Sahano, H., Traunecker, R. & Tonegawa, S. *Nature* 290, 565 (1981)
- 4 Early, P. *et al Cell* 19, 981 (1980)
- 5 Sakano, H. *et al Nature* 286, 676 (1980)
- 6 Burrows, P., Lejeune, M. & Kearney, J. F. *Nature* 280, 838 (1979)
- 7 Riley, S. C., Brock, E. J. & Kuehl, W. M. *Nature* 289, 804 (1981)
- 8 Gearhart, P., Johnson, N. D., Douglas, R. & Hood, L. *Nature* in the press
- 9 Bothwell, A. L. M. *et al Cell* in the press

The expanding Earth

from A. D. Stewart

TWENTY-FIVE years ago the University of Tasmania hosted a conference on continental drift organized by Professor S. W. Carey, at which he presented his now famous review documenting the detailed geological evidence for continental break-up and separation during the Mesozoic and Cenozoic, with explanations for major Earth structures like oroclines (hair-pin bends in mountain belts) and sphenochasms (wedge-shaped oceanic areas like the Bay of Biscay) thrown in for good measure Most of Carey's views on inter- and intracontinental motions are now, of course, generally accepted but the basic mechanism he postulated — Earth expansion sufficient to have generated all ocean basins since the Palaeozoic — engenders in most geologists the same feelings of incredulity that continental drift did 25 years ago

February 1981 saw the sequel to the Tasmanian meeting — the Expanding Earth Symposium at the University of Sydney, again organized by Carey — at which two new lines of evidence supporting major Earth expansion were discussed The first was the work of Andy Glikson

(*Geology* 7, 449, 1979) which highlights the lack of geochemical evidence for oceanic crust, or indeed any mantle-derived crust, surviving from the Proterozoic, and the palaeomagnetic and stratigraphical evidence for continental integrity over this period This was backed up by Keith Crook's analysis of the global abundance of ophiolites and turbidites, both generally accepted indicators of oceanic crust Their abundance decreases exponentially with increasing age back to the end of the Proterozoic (~ 10⁹ years ago) when they disappear completely from the geological record This is not just a sampling effect because turbidite sediments are well known from the Archaean (> 2.5 × 10⁹ yr) Crook's interpretation of these data is similar to Glikson's — there were no ocean basins in the Proterozoic because the Earth had only about 60 per cent of its present radius

The second new line of evidence favouring expansion comes from palaeomagnetism Embleton and co-workers (*J Geophys* 49, 20, 1981) showed that for quite long periods during the Proterozoic, continental blocks such as North America and Australia in their present relative positions had common polar wandering paths They interpret this surprising result,

A. D. Stewart is in the Department of Geology at Reading University

Plesiosaurs — rigging and ballasting

from Michael A Taylor

ALTHOUGH plesiosaurs were among the dominant marine reptiles of the Mesozoic era, it was difficult to see how they hunted their prey for they had large air-filled lungs and would have been too buoyant to dive underwater easily. In a recent paper D G Darby and R.W Ojakangas (*J. Paleont* 54, 548; 1980) suggest that plesiosaurs neutralised their upward buoyancy by swallowing masses of pebbles and retaining them in their stomachs in much the same way as modern crocodiles do.

At one time it was thought that the stomach stones of crocodiles and plesiosaurs helped in grinding food in the stomach. Darby and Ojakangas draw attention to the work of H B Cott (*Zool Soc London Trans* 29, 211, 1961) who found fragile but unpulverised bones and shells of prey along with stones in crocodile stomachs. He went on to show that crocodiles swallow stones to use as ballast so they can lie hidden under the river surface with only the eyes and nostrils protruding above water, until some prey animal comes close enough to be taken. Crocodiles can also lie on the river bottom lodged against the current. The mass of pebbles lying forward in the belly stabilises the animal in roll about the longitudinal axis, like the keel of a boat, and corrects its tailheaviness allowing it to lie horizontal in the water. The weight of the pebbles also helps the crocodile to drag large prey underwater so that they can be drowned. Darby and Ojakangas point out that stones and unpulverised prey have also been seen in a fossil plesiosaur (B Brown *Science* 20, 184; 1904) and so conclude that plesiosaurs also used their stomach stones for ballast.

Their conclusion has implications for our ideas of the hunting strategies of plesiosaurs. If they neutralised virtually all their upward buoyancy, they could stay dived without using much limb action, in contrast to the extinct ichthyosaurs and modern whales which have no stomach stones, but are always on the move when looking for prey and can produce a compensating downthrust with their fins. One group of plesiosaurs, the plesiosauroids, had small heads on exceedingly long necks which gave the animals a most unstreamlined shape and perhaps made them too slow to pursue prey in open water. One wonders if they lurked in ambush, or searched slowly over the bottom ready to strike out their necks at any passing prey. The other main group of plesiosaurs, the pliosauroids, were large carnivores perhaps comparable with the modern killer and sperm whales. They may

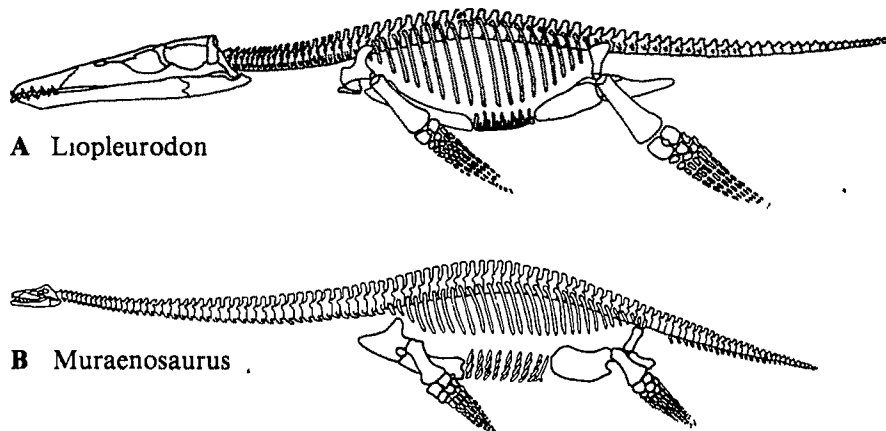


Fig 1 Representatives of the two main groups of plesiosaurs A, the pliosauroid *Liopleurodon* (~11m), B, the plesiosauroid *Muraenosaurus* (~5m) (after Robinson 1975)

have hid unmoving till prey came near enough to be run down in a short sprint, but they were quite streamlined with their large heads and short necks and might equally have been open-water cruising predators. Confirmation of these hypotheses must wait for an analysis of the functional anatomy of the head and neck of these forms (see Fig 1).

Darby and Ojakangas also recovered stomach stones from a plesiosaur found in the Upper Cretaceous of Montana and concluded from an analysis of the shapes of the pebbles that they were collected in an estuary rather than from a shingle beach. It does indeed seem likely that plesiosaurs would forage in the sheltered and productive tidal waters of an estuary rather than off an exposed and comparatively foodless beach. The plesiosaur may, however, have tended to select certain shapes of pebble, so that its stomach stones may not necessarily be a reliable guide to their source.

The structure and function of the distinctively domed body, sprouting four large broad-bladed limbs, has itself been

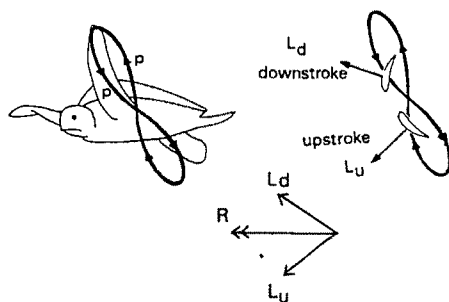


Fig 2 Plesiosaur limbs worked like those of turtles, shown here. The hydrofoil surfaces swept up and down producing lift forces L_u and L_d which averaged out over the cycle to a forwardly directed propulsive resultant force R as shown in the parallelogram of forces (after Robinson 1975)

the subject of an elegant analysis in recent years (*N. Jb. Geol. Palaont. Abh.* 149, 286; 1975 and 153, 86, 1977).

Jane Ann Robinson distinguishes between rowing and underwater 'flying' as the two possible modes of action of the limbs. These two modes differ radically in limb form and action. In the power stroke of a rowing action, the limb sweeps backwards, while its vertically aligned blade bites into the water and levers the animal forwards. The limb then returns forward to start another power stroke. A limb which is adapted to rowing is therefore rigid and oar-shaped, broadening at the end, and moves in the horizontal plane. The limb of a plesiosaur is of quite different design and has the characteristic adaptations to underwater flight. The limb is flat with an aerofoil (or hydrofoil) cross-section so that when it moves through the water with the plane of the limb edge-on to the direction of motion through the water it generates a lift force. This lift force is perpendicular to the plane of the limb and its direction of motion through the water. The plesiosaur limb has therefore to move in the vertical plane to produce the horizontally and forwardly directed propulsive force, unlike a rowing limb which moves in the horizontal plane (see Fig 2). The plesiosaur limb tapers towards its tip to reduce drag, unlike the oar-like shape of the rowing limb. The plesiosaur limb has just enough flexibility to allow trimming and to damp out drag-producing eddies, which is, again, quite unlike the rowing limb which is rigid and relies on its production of drag to bite into the water and thus gain leverage and efficacy. Indeed, the limbs of plesiosaurs share these features with the forelimbs of marine turtles, seals and penguins, which all act as hydrofoils. Robinson has therefore been able to base her reconstruction of plesiosaur limb action

Michael A Taylor is in the University Museum, Oxford

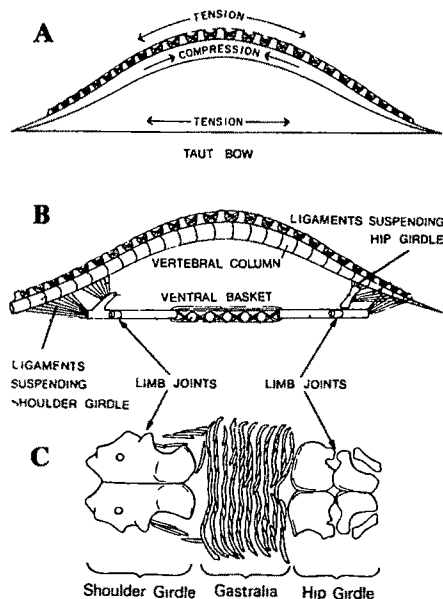


Fig. 3 The shoulder and hip girdles were suspended within the ventral basket which was held under the vertebral column like the string in an archer's bow. This elastic suspension could absorb the up-and-down motion of the limb girdles at each beat of the limbs, while transmitting the propulsive force to the body. A and B, side views of the body, C is a plan view of the ventral basket (after Robinson 1977)

partly on the action of the forelimbs in these modern animals

The distinctive shoulder and hip girdles of plesiosaurs were broad, flat bony plates with limb joints on opposite edges. The main propulsive muscle masses lay on top and under the girdles, and their tendons attached to the dorsal and ventral surfaces of the limbs and swept them powerfully upwards and downwards. Supplementary muscles controlled the trim of the wings so that on both up- and down-stroke the hydrofoil surfaces of the limbs, moving in the vertical plane, generated a resultant 'time-averaged' lift force forwards. Effectively, therefore, each pair of limbs mounted on its girdle was a self-contained power unit producing a forwardly directed propulsive force.

How were the two 'power modules' supported within the domed body, so as to transmit the propulsive forces? Plesiosaurs had a distinctive 'ventral basket' made up from the two limb girdles bound together by ligaments running onto and between the gastralia, the splint-like bones of the middle abdominal floor. Robinson shows that this acted as a string to the 'archer's bow' formed by the vertebral column. The contraction of the muscles between the dorsal processes of the vertebrae bent the ends of the column upwards and pulled taut the ventral basket suspended from the vertebral column by its ligamentous attachments. With the ribs, this formed a remarkable suspension system restraining the girdles while allowing some elasticity to absorb the up-and-down motion of the girdles with each beat of the limbs (see Fig 3). The ability to vary independently the

action of each limb would be most useful, as it made possible rapid changes of direction in pursuit of prey, or when diving from a position at the surface. The propulsive forces of the four limbs would then be asymmetric in magnitude and direction, but they could still be sustained by the girdle support system. One can also conclude that the plesiosaurs could move and breed on land, *pace* Robinson, as the arched vertebral column prevented the collapse of the lungs. Ichthyosaurs and modern whales cannot move on land and therefore give birth in the water to live young; certainly, there are specimens of ichthyosaurs — but not plesiosaurs — with unborn young still in the body cavity. Plesiosaurs may have given birth to live young or laid eggs in pits scraped out with their tails, like modern turtles.

The plesiosaur tail was not very long and part of it was taken up in the 'archer's bow' system. The tail vertebrae bear no evidence of the attachment of a substantial propulsive tailfin — which would not be needed, with four powerful limbs — but plesiosaurs did have small vertical tailfins, presumably used for stabilising the animal directionally and for controlling direction of motion.

The limb and body plan of plesiosaurs, so uniform within the group, is radically different in form and function from the original terrestrial reptilian plan, which was adapted to support the body against the vertical pull of gravity while walking on land. The nothosaurs of the Triassic were small marine carnivores not very different in body or limbs from the terrestrial reptiles. However, the rare nothosaur *Pistosaurus* is morphologically intermediate between other nothosaurs and plesiosaurs and is presumably closely related to the ancestors of plesiosaurs. Nobody has as yet done an analysis of these animals comparable to Robinson's work and so we do not know when the characteristic locomotor pattern of plesiosaurs actually evolved. The plesiosaurs diverged rapidly into their variously adapted lineages at the very start of the Jurassic, in a good example of the macroevolutionary concept of 'adaptive radiation'. Did this event coincide with, and was it caused by the full development of the plesiosaur locomotor plan, or was there some other factor, such as the mass extinction of potential competitors, which allowed the plesiosaurs to exploit the potential of a previously evolved locomotor plan? □

Strategies of viral oncogenesis

from John Wyke

STUDIES on the molecular mechanisms of neoplasia by oncogenic RNA viruses (retroviruses) have in the past concentrated on agents that transform cells in tissue culture. These viruses contain genetic information, oncogenes, that seem essential for the induction and/or maintenance of cell transformation and that are distinct from the genes required for virus replication. Viral oncogenes are believed to be derived from the homologous sequences, present in normal vertebrate cells, that were presumably acquired by the viruses during their evolution. Their reintroduction into the cell under the control of viral transcriptional promoters leads to aberrant oncogene expression and resultant cell neoplasia.

Attention has recently turned to the many retroviruses that do not transform cells in available tissue culture systems but nonetheless induce tumours in animals. These viruses seem to possess only those genes needed for their replication. Without specific oncogenes, how do they induce neoplasia? There are, *a priori*, several possibilities: (1) expression of viral replicative genes themselves may directly or indirectly alter cell growth, (2) since the DNA copy of the retroviral genome, the

provirus, integrates into the cell DNA, it may act as an insertional mutagen, its presence disrupting normal cell gene expression, or (3) the inserted provirus may subvert functions more specifically, by placing particular cell genes under the influence of viral promoters. Two recent papers^{1,2} provide evidence for the last possibility from the avian lymphoid leukaemia viruses (ALV).

Both studies have taken a similar approach, based on recent understanding of the nuances of provirus structure and function that are summarized in the diagram. The coding sequences in the viral RNA (*gag-pol-env* in diagram) are flanked by 5' and 3' unique segments, U5 and U3, and a short terminal redundancy, R. These are thought to be involved in synthesis of viral DNA which eventually becomes integrated as a linear provirus (shown in diagram). The proviral DNA is longer than the parent RNA, U3 being repeated at the 5' end with U5 duplicated at the 3' end of the molecule. The viral genes are thus bounded at each end of the provirus by a long terminal repeat (LTR) sequence, comprising the regions U3-R-U5. Sequencing and *in vitro* transcription studies indicate that U3 contains a putative promoter sequence, and the promoter in the left-hand LTR is thought to direct transcription of progeny viral RNA. Because of the redundancy of the LTR

John Wyke is in the Tumour Virology Laboratory, Imperial Cancer Research Fund Laboratories, London

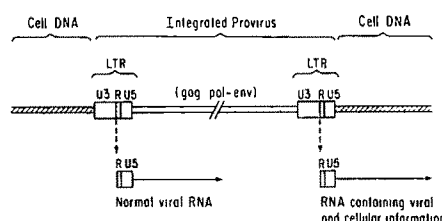
there is also a promoter at the right-hand end of the integrated provirus that seems to have no role in the viral life cycle. However, preliminary evidence indicated that it might also be active and occasionally direct transcription into neighbouring host DNA sequences³.

Starting from this point, Payne *et al*¹ and Neel *et al*² investigated integrated provirus structure and function in a number of advanced lymphomas induced in chickens by several different ALV. They drew the following conclusions: (1) The cells from individual tumours showed provirus integration at identifiable sites. This suggested that the tumours arose largely, if not entirely, from a single infected cell, implying that tumour induction is a rare event. Although many tumours contained several proviruses, those which carried only a single provirus were most useful for further interpretation. (2) Tumours from different birds often showed similarities in the sites of provirus integration. This evidence for tumour-specific integration sites was not strong but was significant in view of the many previous studies showing that proviruses are able to integrate at numerous, perhaps even random, locations. (3) Many tumours contained incomplete proviruses even when only one provirus was present (these latter cells did not contain normal viral RNA species). Thus viral replicative functions are not needed for the maintenance of neoplasia, indeed, immune mechanisms directed against viral antigens might favour the survival of cells whose proviruses do not direct the synthesis of viral proteins. (4) All tumours contained at least one LTR. (5) All tumours contained virus-related RNAs that possessed too little virus-specific material to be revealed by a probe representing the whole provirus but could be detected with a probe for the R-U5 sequences. These RNAs, which fall into only two or three size classes, were expressed at a high level and are thought to represent transcription initiated in a viral LTR (and hence containing U5 sequences) and continued into downstream host sequences (see diagram).

On the basis of these findings Payne *et al* and Neel *et al* propose that ALV cause tumours by inserting a viral LTR, containing a promoter, adjacent to cell sequences that are then expressed under viral control. The evidence for proviral integration in a specific region of the cell DNA and the limited number of size classes of RNA found among virus-related transcripts suggested that viral promoter insertion might lead to neoplasia only if it occurred near one or a few appropriate cell genes.

This possibility was pursued by Hayward *et al* (see *Nature* April 9, 290; 475, 1981), who reasoned that ALV might 'switch on' the cell homologues of the known oncogenes of transforming retroviruses. Accordingly, they screened lymphoma

RNA with cDNA probes specific for five different viral oncogenes. In general, there was no enhanced expression of sequences homologous to four of these, but in 12 out of 14 tumours Hayward *et al* found increased expression of RNA homologous to the oncogene of the acute leukaemia virus MC29 (an oncogene variously called *mac* or *myc*). The *myc*-specific RNA, which was not expressed above background levels in non-tumour tissue from



Structure and transcriptional products of the integrated ALV provirus

the same diseased birds, co-migrated with the same limited size classes of RNA that are detected in tumours with a probe for R-U5 viral sequences, suggesting that the *myc* sequences in the RNA were covalently linked to viral U5 sequences. This conclusion was strengthened by restriction enzyme digests of tumour DNA which showed that fragments containing proviral and flanking cell DNA were detectable with *myc*-specific probes, indicating that the provirus is integrated near the cell homologue of *myc* (*c-myc*). They concluded that in the bulk of ALV-induced lymphomas a viral LTR integrates near *c-myc*, enhances its expression, and this leads to neoplasia.

The data in favour of this promoter insertion mechanism of oncogenesis are attractive but there are a number of inconsistencies which must be clarified before this concept can be accepted. (1) In a minority of lymphomas there is no evidence for enhanced *c-myc* expression. (2) MC29 causes carcinomas and myelocytomas, not the B-cell lymphomas associated with ALV infection. In MC29 *myc* is expressed as a fusion protein, containing additional amino acid sequences encoded by a viral structural gene. Is this enough to explain the disparate pathogenesis? (3) Payne *et al*¹ and Cooper and Neiman (quoted in Hayward *et al*⁴) also find evidence for ALV insertion near *c-myc* but in some cases the provirus integrates downstream of the cell gene and it is not clear how the viral promoters directly determine *c-myc* expression (Payne *et al*¹ and P. Neiman, personal communication). The possibility of the virus acting as an insertional mutagen cannot be dismissed, particularly in view of the structural similarities between the integrated provirus and transposable genetic elements (discussed recently in *News and Views*⁵ and see Neiman *et al*⁶). (4) Cooper and Neiman recently showed that DNA from ALV-induced lymphomas

can transform recipient mouse cells⁷. This transforming DNA contains neither viral structural genes nor viral LTR regions and they have now shown (quoted in Hayward *et al*⁴) that the transforming DNA also lacks *c-myc* sequences.

Some resolution of these discrepancies can be attempted by postulating that *c-myc* functioning is needed for initiation but not for maintenance of lymphoid leukaemia. Provirus integration may stimulate *c-myc* expression (or perhaps maintain expression if the gene is already active) but this may serve simply to render the cells susceptible to further oncogenic change. A similar mechanism has been proposed for Burkitt's lymphoma in man, in which Epstein-Barr virus is thought to stimulate proliferation of B cells that are targets for additional neoplastic events, possibly involving chromosomal alteration. This view is supported by the natural history of avian leukaemia although lymphomas are not seen until four to six months after infection, ALV-induced nodules can be seen in the lymphoid organ, the bursa of Fabricius, within six weeks. The status of ALV and *c-myc* expression in these nodules has yet to be fully investigated. The nodules seem to be clonal, but only a subset of the cells they contain may progress to become metastatic lymphoma cells⁸. It is mainly these lymphoma cells that were studied by Payne, Neel, Hayward and co-workers^{1,2,4}, and the failure to detect either virus or *c-myc* expression in some of the cells, and in all transforming DNA⁷, may reflect tumour progression unrelated to the events initiated by ALV infection.

Whatever the role of *c-myc* in neoplasia, the finding that ALV infection is regularly accompanied by the expression of the cell homologue of a viral oncogene forges a link between mechanisms of oncogenesis by transforming and non-transforming retroviruses. The concept that cell genes can be neoplastic when controlled by a virus promoter has been seminal in recent investigations on the genetic changes that occur after chemical carcinogenesis (discussed in a recent *News and Views*⁹). The possibility that non-transforming retroviruses activate similar genes adds coherence to oncogene hypotheses of tumour formation and provides another refined system in which to investigate such ideas. ALV oncogenesis will clearly be studied in yet more detail and parallel investigations are doubtless being pursued with other retroviruses that do not carry their own oncogenes. □

1. Payne, G. S. *et al* *Cell* 23, 311 (1981).
2. Neel, B. G. *et al* *Cell* 23, 323 (1981).
3. Quintrell, N. *et al* *J. molec. Biol.* 142, 363 (1980).
4. Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* 290, 475 (1981).
5. Flavell, A. *Nature* 289, 10 (1981).
6. Neiman, P. E., Beemon, K. & Luce, J. A. *Proc. natn. Acad. Sci. U.S.A.* 78 (in the press).
7. Cooper, G. M. & Neiman, P. E. *Nature* 287, 656 (1980).
8. Neiman, P. E., Payne, L. N. & Weiss, R. A. *J. Virol.* 34, 178 (1980).
9. Rigby, P. W. J. *Nature* 290, 186 (1981).

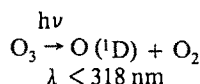
Measuring atmospheric gases and aerosols

from Stuart A Penkett

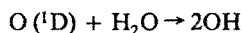
APART from the formation of ions, there are two major photochemical processes which have been identified in the atmosphere. One results in the formation of a layer of ozone in the lower stratosphere and the other gives rise to a general oxidation process in the troposphere. The theory accounting for this second process, although put forward more than ten years ago¹, is not yet fully confirmed. It predicts that many chemical species such as hydrocarbons, halocarbons, sulphur and nitrogen compounds, are oxidized by hydroxyl radicals in the troposphere and converted to more soluble species which can be removed from the atmosphere by rain. This prevents many compounds, which are emitted both naturally and as pollutants, from reaching the stratosphere and interfering with the equilibrium of the ozone layer. The experiment entitled GAMETAG (Global Atmospheric Measurements on Tropospheric Aerosols and Gases) was an attempt by a group of American scientists to provide some evidence on the workings of this process. A series of twelve papers describing the work which took place in 1977 and 1978 has appeared in a recent issue of the *Journal of Geophysical Research*².

By the standards of atmospheric chemistry the experiment was a big one, based around the facilities provided by the National Center for Atmospheric Research at Boulder, Colorado (NCAR). It involved about forty scientists under the general guidance of Douglas Davis, C S Kiang, Paul Crutzen and Robert Duce and utilised a good proportion of the NCAR facilities for two years, collecting data over a flight track extending from Northern Alaska, down the North American continent and across the Pacific to New Zealand.

The most ambitious part of the experiment involved attempts to measure the hydroxyl radical concentrations both in the boundary layer and in the free troposphere above it. Hydroxyl radicals are formed by photolysis of ozone with light at wavelengths below 318 nm, which is about 30 nm above the minimum wavelength of solar radiation that penetrates the troposphere



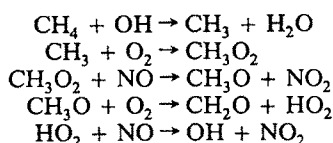
The excited oxygen atom which is formed reacts with the large amounts of water vapour present to give rise to hydroxyl radicals



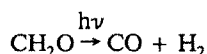
The concentration in clean air has been estimated to be about 5×10^5 molecules per cm^3 and various measurements in polluted air suggest a concentration range of 5×10^6 to 5×10^7 molecules per cm^3 (see refs 3,4). The best ground-based measurements have used a long pathlength UV absorption instrument which is not yet capable of measuring levels below 1×10^6 molecules per cm^3 . The fluorescence instrument used on GAMETAG was unlikely to improve on this, so although disappointing, it is perhaps not surprising that no data on the OH radical concentrations are presented.

Measurements were made on the global variation of the ozone and water vapour concentrations and of the photoproduction rate of singlet oxygen atoms, these enable the radical concentrations to be calculated by two-dimensional models of atmospheric chemistry which are currently being developed. In a world where measurements are not possible, calculations will have to suffice.

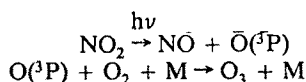
The radicals react with atmospheric trace gases such as methane in a complex chain reaction



The formaldehyde is photolysed to give carbon monoxide and hydrogen



and NO_2 is photolysed to give nitric oxide and oxygen atoms



The ozone which is formed in this way, and also ozone formed from CO oxidation, represents an alternative source to transport from the stratosphere⁵. The experiment showed that distinct ozone layers do exist in the troposphere, but in those observed over the ocean an inverse correlation between the ozone and water vapour concentrations indicated a stratospheric source. There was also no positive correlation between the ozone and carbon monoxide concentrations. The reason is almost certainly lack of sufficient nitric oxide to sustain the chain reaction. Recent measurements⁶ over the Pacific Ocean confirm this and the conclusion

must be that the extent of the atmospheric oxidation process over the ocean is very limited and far from producing ozone, photochemical reactions act as a sink.

Measurements of sulphur and halogen gases and many aerosol components both in and above the atmospheric boundary layer were also made. The sulphur measurements were particularly interesting since they suggest that sulphur dioxide is formed in the free troposphere above the boundary layer. There is also evidence that much of the sulphate aerosol observed was formed by a gas to particle conversion process occurring within the atmosphere. Atmospheric sulphur chemistry has been a subject of discussion for three decades and more. Much of this has focused on the natural sources of sulphur, whether they are present as hydrogen sulphide or dimethyl sulphide. Recently two more gases, carbonyl sulphide and carbon disulphide, have been shown to be present in the atmosphere⁷ and it has been suggested that these are the main precursors of sulphur dioxide and the sulphate aerosol found in the remote troposphere⁸. The GAMETAG experiment has certainly proved beyond doubt that carbonyl sulphide is the most abundant sulphur-containing trace gas in the atmosphere. However, it is not clear yet how quickly it is oxidized^{9,10}.

GAMETAG showed very little evidence for the presence of long-range transport of pollutants from the continents over the ocean. This was true of gases such as carbon monoxide and sulphur dioxide and for aerosol species such as sulphate and nitrate. It is also clear that the ocean is the major source of methyl chloride, which is the only naturally produced halogen-containing gas to enter the stratosphere in large quantities.

In conclusion, GAMETAG provided some interesting data on the concentration distribution of stable trace gases and aerosols and the influence of pollutants on the composition of the remote troposphere. However, it failed to provide positive evidence for the existence of a large amount of chemical activity in the troposphere. Most probably this is associated with low levels of nitric oxide over the oceans, which suppresses the level of hydroxyl radicals.

Stuart A Penkett is in the Environmental and Medical Sciences Division, AERE, Harwell, UK

- 1 Levy, H. *Science* **173**, 141 (1971)
- 2 *J. geophys. Res.* **85** No C12, 7285 (1980)
- 3 Perner, D. *et al. Geophys. Res. Lett.* **3**, 466 (1976)
- 4 Wang, C C. *et al. Science* **189**, 79 (1975)
- 5 Fishman, J. and Crutzen, P J. *Nature* **274**, 855 (1978)
- 6 McFarland, M. *et al. Geophys. Res. Lett.* **6**, 605 (1979)
- 7 Sandalls, F J and Penkett, S A. *Atmos. Environ.* **11**, 197 (1977)
- 8 Logan, J A. *et al. Nature* **281**, 185 (1979)
- 9 Kurylo, M J. *Chem. Phys. Lett.* **58**, 238 (1978)
- 10 Cox, R A. and Sheppard, D. *Nature* **284**, 330 (1980)

The continents are likely to be the major source of nitrogen oxides and these are the places to look for evidence of atmospheric reactions. There is certainly a case for an experiment of this type to continue if only to confirm that many compounds released from modern industry are destroyed before they can reach the ozone layer in the

stratosphere. It should concentrate on unpolluted continents in tropical latitudes and it should include measurements of transient products such as aldehydes and nitro compounds. These are easier to measure than radicals but provide the sort of information which is necessary to test theories of atmospheric oxidation. □

Cracks in rocks under stress

from Barry Atkinson

MICROCRACKS influence many of the physical properties of rocks. Strength, transport, elastic and inelastic characteristics are all highly dependent upon the size, shape, number and distribution of microcracks. These physical properties change dramatically as crack characteristics are modified under stress.

In earthquake prediction, the exploitation of oil, gas and water bearing formations, nuclear waste disposal, and heat extraction and storage, the prediction of *in-situ* physical properties from models of rock microcracks is of great importance but has, so far, been rather unsuccessful. The basic problem is that, despite many laboratory studies, we have only a rudimentary idea of the actual modifications that cracks undergo in stressed rocks. M. L. Batzle, G. Simmons and R. W. Siegfried of MIT have recently attempted to overcome this problem by making direct observations of crack closure in granite and diabase under uniaxial compressive stress (*J geophys Res* **85**; 7072, 1980).

The MIT group placed a small hydraulic press inside a conventional scanning electron microscope (SEM) so that they could subject cylinders about one cm in diameter and two cm long to uniaxial stresses of up to 30 MPa. A portion of the lateral side of the core was ground flat and bombarded with ionized argon to remove the damaged surface before examination in the SEM. Stress-strain measurements were also made on similar cores in the press but outside of the SEM. Although SEM studies have been made before on specimens subjected to tensile *in situ* stresses, this is the first time that a compressive stress has been so studied.

Westerley granite and Maryland diabase were examined in the 'natural' state and after slowly heat cycling to 500°C and 700°C, respectively, to increase crack porosity. Westerley granite is a typical crystalline rock with initial crack porosity of about 0.2 per cent and a total porosity of about 1 per cent. Maryland diabase is unusual in that it has no measureable crack porosity and a low total porosity of 0.1 per cent. Westerley granite consists of roughly equidimensional feldspar, quartz and

biotite mica grains averaging 800 micrometres in diameter. Maryland diabase consists mainly of equidimensional grains of feldspar and pyroxene averaging 200 micrometres in diameter.

In general, the response of cracks to stress confirmed those predicted from microscopic examinations made at zero stress (Kranz, R. L. *Int J Rock Mech Mining Sci Geomech Abstr* **16**; 37, 1979) and from deformation measurements made under applied stress (Brace, W. F. *et al J geophys Res* **71**, 3939, 1966). Numerous factors influence the behaviour of cracks, including orientation, shape, source, and the proximity of other fractures.

The main conclusion of the recent work, however, is that pre-existing cracks in these rocks behave quite differently to those induced by heat cycling. Pre-existing cracks have walls that are irregular, etched and pitted, and poorly matched. Under stress, closure is incomplete with many portions remaining open and interconnected. The shape, or aspect ratio, continually changes as stress increases. Thermally induced cracks are also irregular but the walls are more perfectly matched and tend to close more completely, unless propped open by very small grains or debris.

Although pre-existing fractures can be re-opened, widened, extended, or closed, depending on a host of interrelated factors, in general those cracks oriented close to the direction of maximum stress tend to open, whereas those oriented nearly perpendicular to the maximum stress tend to close.

New cracks created in the laboratory by stress or temperature usually have the irregular but well matched characteristics of thermal fractures. This is important because the precise nature of the crack response to stress determines the trends in dependent physical properties. New fractures will close easily and hence lead to a large change in physical properties with stress as compared with the more inefficient

closing of pre-existing cracks.

How then do induced cracks attain the characteristics of pre-existing cracks? A key element in this seems to be the extent to which cracks are modified by the circulation of water at high temperatures. Under these conditions grains become rounded and crack walls etched and pitted (Potter, J. M. *Rep LA-7224-T*, Los Alamos Sci. Lab., 1978) and begin to resemble pre-existing cracks.

The fact that almost complete closure of thermal and stress-induced cracks can occur lends support to D. J. Holcomb and J. L. Stevens' assertion that reversible Griffith cracks can explain all of the observed attributes of rock dilatancy (*J geophys Res* **85**; 7101, 1980). They cast doubt on the widely used sliding crack model of dilatancy, in which dilatancy arises from the formation of tensile cracks at the ends of a large number of sliding shear cracks, partly because a shear crack model cannot explain the stress memory rocks exhibit and especially because shear cracks are only rarely observed in some dilatant rocks (Tapponier, D. and Brace, W. F. *Int J Rock Mech Mining Sci Geomech Abstr* **13**; 1976). On the other hand, long, narrow cracks preferentially oriented parallel to the maximum compressive stress are very common in some rocks and these Griffith cracks can apparently explain all of the experimental results on rock dilatancy, provided that the cracks can open and close reversibly under stress.

A problem arises, however, because most studies of stressed cracks in rocks have been made on low-porosity, well-cemented crystalline rocks. Less well-cemented, high-porosity sedimentary rocks with more point contacts and fewer interlocking grains have quite different initial preferred orientations and distribution of microcracks. Recent experimental studies have shown that the differences in peak strength, the localization of deformation to give a main fault, and some water-weakening effects can be qualitatively explained in terms of the initial microstructure (Hadizadeh, J., Thesis, University of London, 1980).

Two limitations of the MIT study are that we are not sure their observations are typical of the bulk of their specimens. They examined sections that were within 1 mm of the surface, and the role of stress corrosion crack growth was not investigated in the high-vacuum environment of the SEM. *In-situ* tensile tests, however, show that stress corrosion can be a significant factor (Swan, G. *Res Rep*, Tulea 1980 01, University of Lulea).

It would now be valuable to extend *in-situ* observations to higher stresses and temperatures and more complex polyaxial stress states. At present this is not possible and one is forced to examine at zero stress and ambient conditions specimens that have been retrieved from more exotic conditions. □

Barry Atkinson is NERC Special Research Fellow at the Geology Department, Imperial College, London.

SPRING BOOKS SUPPLEMENT

Reviewers without pride or prejudice

What are book reviews for? The question is not entirely trivial. The pages that follow are filled with the comments of a great many serious people who have taken the trouble to read some book or books, to reflect on the contents thereof and then to tell the rest of the world what they think of them. Why, it will be asked, should they bother? Why should an army of people who are (or who should be) disinterested take this trouble for the sake of others — the people who lie somewhere between those who would certainly buy the book concerned, perhaps because it represents a special interest of theirs, and those who would not willingly be seen in public reading the document concerned, and who would not wish to do so in private either? Pecuniary reward is not the explanation. The simple answer is that book reviewers pursue one of the most honourable of intellectual crafts. The ideal reviewer is a self-effacing person — the kind of person who is sufficiently knowledgeable to have written the book he is charged with off his own bat, who might well indeed have done so if he were not so busy, but who has nevertheless been willing to quarry time from a busy timetable to inform and instruct potential readers of a new apparition in the literature. The scientific community, any intellectual community, has reason to be grateful to reviewers.

But why are books so especially deserving of attention? What is it about a pair of hard covers that gives a book precedence over, say, an article in a scientific or other journal? Rarity is only part of the explanation. Some journals, *Nature* among them, do indeed try to pick out and draw to the attention of their readers particularly interesting articles that appear in the ordinary course of communication. The volume of the scientific literature is of course too great for this task to be attempted (or at least accomplished) systematically. There is, however, another sense in which a book commands (and deserves) to be dealt with differently from the report of some piece of research, whatever the excitement engendered by the results. For a book, or at least a good book, is necessarily a somewhat personal declaration. So much can be told from a comparison of the style of run-of-the-mill scientific articles and books by the same authors. The authors of articles run to the passive, and to opening sentences such as "There is now much evidence to suggest that ...". The authors of books write in a different frame of mind. They are, from time to time, prepared to utter sentences beginning "I believe that ...". More than this, books differ from articles in the journals in their objectives. Both kinds of contributions to the literature have a beginning and a middle, but only of books (or good books) is it required that there should also be a conclusion of some weight and finality. The authors of articles in the journals are allowed to finish with a statement to the effect that they are entirely at a loss to understand the observations they report and to declare that they will do their best, in competition with those who are stimulated by reading their contribution to the literature, to find out. Would-be book authors, by contrast, must (or should) know before they put pen to paper what case they seek to make. (Publishers who make books by stringing together a miscellaneous collection of articles, or collections of papers given at some symposium, must by the same token appreciate that these are comparatively dull, not much like books.)

There is another crucial respect in which a book differs from some other contribution to the literature. Although some books may only be ten times longer than, say, some journal articles, the

difference of scale is a quantitative difference big enough to imply a qualitative difference. As the world knows, book authors are repeatedly perplexed to know whether the point they are about to make in chapter umpteen has already been made in chapter one. Put simply, those who set out to write books, and actually finish them, command a certain respect, even from those who may disagree with the conclusions.

Good reviewers appreciate these truths, and are (or should be) compassionate people. If somebody writes a routine account of, say, the inorganic chemistry of selenium, and if the arrangement of the chapters is not precisely what the reviewer would have chosen, that does not amount to an excuse for a fierce attack on the author and his good sense and a warning that students everywhere should avoid the book like the plague. In such fields it is well known that the unkindest course a reviewer can follow is to say nothing at all: then potential purchasers will have to learn of the book by chance, while the book's publishers will be at their wits' end, seeking other ways of putting news of the book's appearance into circulation. In the last resort, after all, the reviewer's task is to tell potential readers what a book seeks to accomplish and to say whether the author has done what he set out to do. Part of the book reviewer's licence is that he may, in passing, add a few reflections of his own on the good sense of the author's goal. So most book reviews should be essays in sympathetic understanding, marked sometimes with sorrow, rarely with anger. And so they usually are.

There are, however, exceptions to this benign rule. For the class of well-conceived books, each with its own sharp conclusion, is bound to include many whose conclusions are at once mistaken, even mischievous, but also likely to be persuasive. This is why the craft of book reviewing frequently runs to polemic. This is why H. J. Eysenck's *The Causes and Effects of Smoking* (Temple Smith, 1980), or Samuel Epstein's *The Politics of Cancer* (see *Nature* 284, 297, 1980) persuaded reviewers to soak their typewriter ribbons in vitriol. The arguments, right or wrong (and probably wrong, or at least incomplete), are such a challenge to accepted doctrine and at the same time so self-consciously challenging that it is helpful to readers to know the strength of opinion on the other side. Polemic is therefore part (but ideally a small part) of book review columns. Professor Dorothy Hodgkin's review of Maurice Goldsmith's book on J. D. Bernal (*Sage, Hutchinson*, 1980, see *Nature* 289, 99, 1981) is more puzzling. Goldsmith set out to write a biography of one of his heroes, and was savaged for his pains by another hero-worshipper, his reviewer. No doubt the book is, as Professor Hodgkin says, "confused and inaccurate". It is also, however, a good book in the sense of being a good read and also (for a hero-worshipper) honest in that it deals with the warts as well as the achievements — sexual proclivities, ambivalence about Lysenko and all that. Maybe Goldsmith's book is "not the book that is needed about Bernal", but it is the only book we have. On this occasion, polemic seems to have been misplaced. In general, the truth about books is that they are milestones for somebody — certainly the author and the publisher — and that there are some occasions on which even bad books are better than no books. Moreover, books differ from other contributions to the literature of scholarship in the diversity of their readership. Their influence on students is important. And books — or good books — can survive from one decade to another.

Advice and dissent in White House science

Harvey Brooks

THE growing literature of science policy generally includes two kinds of books: first-hand accounts of their own experience by principal actors in the processes of high-level scientific decision-making, such as those by James Killian, George B. Kistiakowsky and Herbert York, or analytical discussions by academic scholars based on study of official documents or interviews with the principal actors. Edward Burger attempts to build a bridge between these two genres. A member of the staff of the Office of Science and Technology (OST) from the time of the first Nixon administration to the abolition of the President's Science Advisory Committee (PSAC) and OST in 1973, Dr Burger has built his book firmly on his own first-hand experience as an insider, but at the same time has attempted to bring a historical and analytical perspective to this experience and to put forth some important generalizations.

To the average observer of science policy, the Nixon administration tends to be viewed as the low point in the history of science in the White House. The rather sudden decision to abolish the presidential science advisory apparatus in January 1973 overshadows the many constructive initiatives that arose during this period. As Burger reminds us, this was a time when the effort to "civilianize" OST and PSAC — to bring scientific considerations and the potential of science fully to bear on some of the nation's major sociopolitical issues such as health care, environmental regulation, population policy and urban policy — reached its climax.

His account is notable for its political sophistication and sensitivity to the multiplicity of influences bearing on even the apparently scientific decisions. It is also refreshingly free of self-justification or overstatement of the role of science. The chapters on national health policy, on environment and health, and regulation to protect health and on population and family planning contain a wealth of new material relating to major Nixon administration initiatives in social policy which gradually became diluted by pluralistic political pressures. These chapters are well worth reading even by those who care little about science policy or the role of science in the White House.

The subtitle of the book is unfortunately somewhat ambiguous and fails to do justice to all the themes which Burger develops. Its meaning is never fully explained and conveys implications that are not really supported by Burger's detailed conclusions. It is thus not clear whether he means to imply that having a Science Advisor in the White House is a

Science at the White House: A Political Liability By Edward J. Burger Jr. Pp 208. ISBN 0-8018-2433-8 (Johns Hopkins University Press 1980) \$14.95, £9.

political liability to the President, which tends to tie his hands in negotiating with the many contending constituencies that he has to satisfy or whether having a prominent presence in the White House is a political liability for the scientific community which compromises its long-term credibility before the public. Perhaps both statements are true, but I do not believe that the implication that science should be taken out of the White House is either intended by Burger or justified by the case histories which he analyses so perceptively.

In fact one could argue that science in the White House enjoyed some of its finest hours during the period covered by this account, if retaining the intellectual honesty of science in the policy process is used as a criterion. The report on *Chemicals and Health*, published by the National Science Foundation after the demise of PSAC, is still one of the most thorough and balanced discussions of this controversial topic. Issued in a climate of overwhelming environmental activism, its sober and balanced approach was little appreciated or heeded until much later. Had it been heeded, some of the current over-reaction against the earlier excesses of the environmental movement might have been avoided. The Toxic Substances Control Act (1976) derived from a concept that originated from the staff of the Science Advisor. The OST staff, and Burger in particular, played a central role in shaping the recommendations of the Health Policy Review Group (1971) and the subsequent Presidential Message on Health. The analysis of the Review Group stands up very well today, though it had little political effect at the time. The OST also played an important role in following up the report of the Population Commission (1972), although here again the analysis was ahead of its time in political terms. The Science Advisor played a reluctant but sensible role in devising the Quality of Life Review Process for considering the social and economic impacts of regulatory decisions, anticipating the regulatory review process established in the Carter administration. The Science Advisor chaired a special committee created by the President to study the rather grandiose and politically motivated proposal for a New Technology Opportunity Program in 1971.

The recitation of these examples, however, poses the dilemma suggested in the title of the book. In referring to these

activities favourably I am making an intellectual — not a political — judgement, because in most instances the input from the OST, as Burger suggests, must be rated a political failure. The intellectual conclusions were mostly out of tune with the political climate and the President's political goals — the classic problem of "speaking truth to power". The report on *Chemicals and Health* took too long to complete and appeared too late to have any significant impact on policy, and its balanced approach was out of tune with the public's mood. By the time the hard hitting and clever recommendations of the Health Policy Review Group had been distilled into a presidential message, much of the bite had disappeared, and health as a salient political issue had faded into the background. Much the same fate befell the report of the Population Commission. The task force to follow up the report completed its work at a time when political concern over US population growth had faded with the dramatic decline in fertility in the early 1970s. Although the task force called for a presidential statement to endorse the Commission's principal recommendation for a national population stabilization policy, the chairman privately sabotaged this recommendation. The technology initiatives programme, which was launched with great political fanfare, quietly foundered when Dr Edward David's standing committee discovered there was much less there than met the eye, and was unable to identify any dramatic government initiatives that stood up to scientific scrutiny.

We thus end with a picture in which strong recommendations based on scientific input failed politically because they ran counter to the prevailing climate or antagonized strong political interests, on the one hand, or in which careful analysis debunked political initiatives on which the President had set great store, on the other. Are the Science Advisor and his staff to be praised for "speaking the truth to power", for calling the shots as they saw them and for thoughtful and forward-looking policy analysis? Or are they to be criticized for insufficient awareness of political realities and of the sensitivities of important political constituencies to which the President must respond? Is the performance of science in the Nixon White House to be viewed as a series of honourable failures derived from commendable devotion to intellectual honesty and objectivity? Or is it to be condemned as the mark of politically naive amateurs incapable of appreciating and responding to the legitimate political needs of the chief executive? Burger poses these



Scientists at the White House, 1957–1981. Left to right, top to bottom: James Killian Jr (served under Truman, 1957–1959); George Kistiakowsky (Eisenhower, 1959–1961); Jerome Wiesner (Kennedy, 1961–1964); Donald Hornig (Johnson, 1964–1969); Lee Du Bridge (Nixon, 1969–1970); Edward David Jr (Nixon, 1970 to 1973 when PSAC and OST were abolished); H. Guyford Stever (Ford, 1976–1977); Frank Press (Carter, 1977–1981). President Reagan has not yet appointed a successor to Press.

questions cogently in each of the examples which he analyses, but he does not really give us the answer. Perhaps there is no single answer and the greatest contribution of the book is revealing the dilemma in realistic detail.

An important theme of this book, set forth in Chapter 2, is the fundamental conflict between rational policy planning and the push and pull of pluralistic political interests. The author presents the executive branch and the scientists as the proponents of such rationality against the pluralism embodied in the legislature. Yet there is a paradox here, which is not fully resolved in the book. It is true that in the efforts to introduce Program Policy Budgeting (PPB), to develop a comprehensive health policy or to formulate a rational and balanced national policy for environmental regulation, Burger's model seems to apply. But he fails to note that in policy for science, as contrasted with science for policy, the roles of the scientists and politicians appear to be reversed. It is Congress that has continually pushed for a comprehensive national policy for science, most notably in the National Science and Technology Policy, Organization and Priorities Act of 1976. Furthermore, this is

not a new phenomenon. In the nineteenth century, Congress was the repeated source of proposals for a cabinet Department of Science. In every case it has been the scientific community and the executive branch that have dragged their feet in regard to the comprehensive planning of science and technology. More recently, OSTP under the Carter administration has attempted to divest itself of most of the planning and systematic reporting functions mandated by the 1976 Act. Whereas Congress has viewed OSTP and the Science Advisor as a kind of general staff for science and technology in the United States, the Science Advisor and most of his constituency in the technical community have viewed his role more modestly and have strongly favoured a pluralistic, decentralized approach to the support of science itself.

It is interesting to speculate on the causes of this paradox of science as the advocate of planning in the spheres of public policy and pluralism in the field of science, while the Congress and the more politically minded Executive take the exactly opposite view. The answer may lie in the internal politics of the scientific community itself. This community is not a monolithic group

but rather congeries of frequently conflicting interests and views. Planning for science, especially by scientists, entails a high political cost within the scientific community, which would rather take its chances on the free play of political advocacy between the different segments of the scientific community and the political process. It would have been desirable if Dr Burger had explained this particular paradox a little further.

Altogether, though, this is a valuable book which should be of interest to a wide public. It provides some fascinating insights into some of the major social policy debates of the early 1970s; it is a contribution to the analysis of the role of White House science advice from the perspective of a staff officer carrying out policy support rather than making policy; and it provides excellent case material for scholars and students in the field of science policy. On top of that it is well written, literate and sophisticated. □

Harvey Brooks is Benjamin Pierce Professor of Technology and Public Policy at Harvard University. He was a member of the President's Science Advisory Committee from 1959 to 1964.

Education and the compleat scientist

J. Z. Young

Teaching and Learning about Science and Society. By John Ziman. Pp.190. ISBN 0-5212-3221-X. (Cambridge University Press: 1980.) £9.50, \$22.50.

MOST scientists are interested in scientific education, if only for the memory of their own debt to it. I imagine that a great many became scientists because they were lucky enough to have an inspiring teacher. My own at Marlborough College was A. G. Lowndes, a remarkable man who trained several Fellows of the Royal Society and a Nobel Prize Winner (Peter Medawar). He inspired us by his own somewhat amateur researches, he showed us endless living things and, by making us work even in our spare time, taught us that all science involves much hard work.

These are perhaps the main items that must be included in any scientific education: it must catch enthusiasm, it must cover a huge mass of facts and it must find out those who can enjoy learning a lot of detail. This is how scientists are made, but in the process they seldom learn much about the position of science in society, still less about its history or philosophy.

It is these defects that are the particular concern of Professor Ziman as indicated by the title of his book. I cannot remember ever hearing them discussed by my teachers either at school or university. Probably many scientists still take up their subject because they are interested and like it, though there is now more concern about its social implications. They are all too liable to fall into what Ziman calls "scientism and technocracy", the uninformed assumption that everything that is done in the name of science is good. How can this restriction of the development of young scientists be avoided? Ziman has tried to encapsulate the problem under the acronym STS — Science, Technology and Society. These themes permeate the political, economic and cultural issues of our times yet "there is no *rationale* of STS as an educational subject", although various courses of study of it have been tried. In trying to provide this rationale Ziman manages to stimulate thoughts about a great many fundamental questions, even beyond those dealing with the educational problem itself.

For instance, one feels that the discussion obviously needs a definition of science, and many times Ziman seems on the point of providing one, but wisely never comes down to calling it a definition. Sometimes it is a question of method: "Science derives its practical power and authority from the rigours of its arguments and the hardness of its facts". Moreover, he repeatedly stresses that the production of scientific knowledge is a social process. Then only a page later we are in a different world of discourse: "There is no single

'scientific' map of reality — or if there were, it would be much too complicated and unwieldy to be grasped or used by anyone". Is this true? I suspect that Ziman doubts it himself. He pursues the question into a discussion of the hierarchy of scientific enquiry and the questions of emergence and reductionism,

the notion that the properties of complex systems such as organisms or molecules can be 'reduced' to the laws satisfied by simpler systems such as cells or atoms — is not only a very dubious philosophy, it is a dangerous folly in science education, where the map appropriate to each level must be taught wholeheartedly according to its own lights

Yet probably most scientists have a hankering for that "folly". And is it indeed so foolish? It is true that chemistry and biology have their own laws but no one can go all the way in either of those sciences unless they know something about physics. However much "autonomy" there may be for each part of science there is still, at least for many of us, a scientific map or model of the world to which we try to refer all events and all knowledge, including the knowledge of ourselves. It seems that the human brain is so constructed that it tries to build all its information and schemes of action around a unified model. Of course this grows gradually in each one of us, centred at first around a parental scheme. The scientific model, if properly learned, can provide the adult substitute, which many people feel they lack in the absence of religion. That is of course not to say that science should be dogmatic or be treated as religion, but that it can provide what Ziman calls "the possibility of discovering order in nature" and to do this is a requirement for human life. We all need some system of order. Incidentally this is a characteristic specific to human beings and not present in our nearest relatives. Ziman actually recognizes this later in his book

when he says that "scientific world pictures" allow "deductions that help us along the way through life".

This is only one of the ends that he believes could be achieved by the proper attention to STS education. He holds that they can only be reached by specifically designed courses for all levels of scientific education. In the later part of his book he discusses how these should be planned for different stages. He is against "General Science that is too sloppy and technical science that is too arbitrary", and he advocates a "valid science". This is easy enough to say until you come to the job of deciding what to put in and what to leave out. Ziman discusses the many ways in which it could be done. For instance, it can be through the social relevance and applications of science, through its possibilities as a vocation for the individual, or its history, or its philosophy or its value in solving world problems. But he does not try to specify the content of particular courses. His aim is to show that the teaching of science at present is grossly defective because it does not properly develop the understanding of the individual either of himself or society, or indeed of the nature of the science he is learning. His plea is that teachers of science should take definite steps to fill this gap by designing what he calls STS courses. He knows that there will be much opposition and anyone who has tried to introduce such courses will have experienced this. But there are many people ready to agree with his plea that a better understanding of the relations of science and society is needed for the proper training of research scientists as well as doctors and other technologists, not to mention civil servants, politicians and the general public. A major value of this interesting book is that it makes you realize how difficult it is to reach agreement either about the nature of science or its proper place in society. □

J. Z. Young is Emeritus Professor of Anatomy at the University of London.

IQ gladiators in separate combat

Stuart Sutherland

Intelligence: The Battle for the Mind. H.J. Eysenck versus Leon Kamin. Pp.192. ISBN hbk 0-333-31279-1/0-471-08884-6; ISBN pbk 0-330-36399-4. (Macmillan Press, London/Wiley, New York/Pan: 1981.) Hbk £12, \$12.95; pbk £2.95.

NATIVISTS have battled with empiricists for many years, but never with such fury as over the inheritability of intelligence. It might be thought that the issue could be resolved by bringing together representatives of the opposing sides and letting them argue out, point by point, the

evidence for the respective roles of inheritance and of environment. *Intelligence: The Battle for the Mind* could have provided just such an opportunity, but unfortunately it is constructed in such a way that the combatants only skirmish and never meet in a decisive battle.

The book contains four sections. In the first, Professor Eysenck sets out his case for supposing that in Western society inheritance accounts for about 80% of the variation in intelligence. Although he does not enumerate them, he deploys 20 different lines of evidence from which he claims this inference can be drawn. In the second

section, Professor Kamin produces his criticisms of some of the evidence for the importance of heredity, but this section is not a reply to Eysenck's initial foray since it was evidently written without Kamin having read it. In consequence, many of Kamin's arguments are beside the point: for example, although Eysenck excludes all reference to Burt's data, now universally acknowledged — thanks to Kamin's brilliant detective work — to have been faked, Kamin devotes a chapter to attacking Burt. In the final two sections, which are much too brief, Eysenck replies to Kamin's set-piece and Kamin to Eysenck's, and there the matter rests.

No agreement is reached on which aspects of the evidence are reliable and which are not, and both authors are guilty of failing to meet the arguments adduced by the other. Thus, in his opening section Eysenck ignores many of the criticisms previously levelled by Kamin. For example, one of the strongest arguments in favour of the role of inheritance is that when monozygotic twins are brought in by foster parents, there is a higher correlation between the IQ of the natural parents and the child than between that of the foster parents and the child. Kamin has put forward two ingenious suggestions to explain this correlation in environmental terms. First, many of the foster children in question spent a year or more of their lives with their natural parents before moving to foster parents. Second, the IQ and the socio-economic status of foster parents is in general high and has much less variance than the IQ of the population at large: this reduction in environmental variance would reduce the effects of environment. One would have liked to have seen Professor Eysenck's reply to these arguments, but it is not to be found in the book. On the other hand, Eysenck points out that there is a higher correlation between the IQs of monozygotic twins reared together, but one looks in vain for Kamin's response to this argument.

Both authors give the impression of being determined to make a case and of selecting data that forward that case. It is a pity that Eysenck did not concentrate on the most solid evidence for inheritability and ignore dubious or unrepeatable findings. For example, he argues that there is a difference in the brain waves (EEGs) of the intelligent and unintelligent and that this difference suggests intelligence is inherited. Not only is the argument fallacious, since there is no reason why brain waves should not be in part determined by environmental factors, but recent attempts to repeat the finding Eysenck uses have failed.

Eysenck points out that Kamin does not have a theory of intelligence. With great ingenuity, Kamin has taken a number of pieces of evidence thought to support the hereditary case and has shown that there are other *possible* explanations: he has not proved that these other explanations are

correct. Moreover, since Kamin treats each piece of evidence in isolation, he is able to use individual arguments that are inconsistent with one another. Thus, as Eysenck notes, Kamin explains the large differences in IQ between dizygotic twins reared *together* by assuming a large difference in the way they are treated within the family. Elsewhere, however, he explains the similarity in the IQs of monozygotic twins brought up *apart* by arguing that each pair is placed in families of a similar socio-economic background. Taken together, Kamin's two arguments imply the absurd conclusion that there is less environmental difference *between* families than *within* a family.

There are a few things on which both authors agree. Moving children from deprived homes to particularly good homes can bring about a shift of up to 20 points in IQ. Although Kamin regards this as compelling evidence for the overriding importance of environment, Eysenck is able to show that it is consistent with his own theory of intelligence. He does not maintain that a large shift in environment cannot change IQ, merely that in conditions as they exist today IQ is more determined by heredity than environment.

Kamin does not help his case by his personal virulence towards Eysenck. The fact that Eysenck has twice made a mistake about the sex of an author whom he cites has surely nothing to do with the inheritance of IQ; and it is merely distracting to the reader for Kamin to stress past errors made by Eysenck, at least some of which (for example, his defence of Burt) Eysenck now acknowledges. Moreover, Kamin's allegation that the aim of the science of genetics is to make "the world comfortably safe for white males" is both vituperative and false — white males were a good deal safer before IQ tests were ever thought of. Eysenck rightly insists that differences in the average IQs of groups have no bearing on how the individual should be treated. It is impossible to predict the individual's IQ from skin colour, sex or social class and in at least some instances the use of objective tests for selection has favoured underprivileged groups. Burt may have been a scientific scoundrel, but his introduction of intelligence tests for secondary education in Britain between the wars doubled the proportion of poor children in secondary schools.

Like all scientific discoveries, the results of work on intelligence testing can be used for good or ill. As Eysenck notes, recent decisions by anti-racists in the United States are likely to be extremely damaging to minority groups. For example, it is folly to insist that equal proportions of blacks and whites should enter training programmes for the educationally subnormal, if, for whatever reason, a higher proportion of blacks than of whites need the help of such training. Eysenck quotes a remark of Dr Johnson's which sums up the position: when asked whether

men or women were more intelligent, he replied: "Which man? Which woman?"

Intelligence: The Battle for the Mind is a wasted opportunity. One feels that had Eysenck concentrated on the most compelling arguments for the role of inheritance, and had he attempted from the outset to meet Kamin's ingenious but *ad hoc* arguments where they could be met and conceded cases where they could not be met, the debate might have been settled. Eysenck might not have proven that IQ is 80% inherited, but he might have established once and for all that there is a strong inherited component.

Neither combatant comments on the scientific importance of intelligence testing. A much more destructive criticism can be made of the whole enterprise than any advanced by Kamin, namely, that it has told us nothing about how the human mind works nor has it given any answer to the really important applied problem — how can we make it work better? □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.

Reality of science

Peter Newmark

An Imagined World. By June Goodfield. Pp.288. ISBN 0-06-011641-2. (Harper & Row: 1981.) \$12.95 US only.

IF A scientist wishes to discover what makes a certain type of cell tick, she may well spend five years studying its behaviour and composition. If a writer wishes to find out what makes that scientist not only tick but chime, by way of making scientific discoveries, she may decide to track the scientist's every thought throughout the five years. That neither venture is guaranteed to succeed is illustrated by this book. The story, however, is not without interest.

In 1975 June Goodfield met Anna Brito (a pseudonym) who was then on sabbatical in New York and, within hours, decided that she fitted the bill of an individual research worker through whom to follow the process of science. Goodfield found Anna Brito "articulate, amusing and somewhat different from the majority of scientists I had met before", and sensed that Brito was at a stage where doors were beginning to open. In one sense, at least, she was right. Within a year of returning from New York to Glasgow University, the doors of New York's Memorial Sloan-Kettering Cancer Center had opened to Anna Brito. From then on Goodfield had ready access to the scientist. So except for the first year and for a period in 1977, when progress is recounted through letters and tape recordings, the tale is told in narrative form.

Through the letters, recordings and narrative emerges a story that illuminates the working life of a scientist in a detailed and open way that has seldom, if ever, been done before. Unlike *The Double Helix*, *An Imagined World* contains little glamour, much toil, many failures, much anguish and no prizes. It is, in a word, realistic. Whether that makes for interesting reading is a matter of taste. Certainly for the working scientist there will be few surprises, but there may be reflections, both pleasant and not so pleasant, to be seen in the mirror of Anna Brito's progress. For the non-scientist who does not realize just how great a travesty of the making of a discovery is the usual account given in a scientific paper, the book is bound to come as an eye-opener.

It would be hard, however, to claim that Anna Brito's experiences were particularly typical. Her first, and arguably most important, discovery certainly came about in an unusual way. She found herself in a London laboratory, unable to speak much English, having neither skills nor ideas to offer and altogether a rather unwanted visitor. Her supervisor, somewhat in desperation, gave her a large collection of sections of mice to examine through a microscope. Against all the odds she noticed something both original and interesting. Her observation amounted to the discovery that T-lymphocytes and B-lymphocytes have predilections for separate zones of tissues.

In the years that followed, Anna Brito pursued her discovery of what she called "ecotaxis" and developed the idea that certain diseases were associated with wrongly located lymphocytes. This condition she called "ecotaxopathy", a term guaranteed to turn purple the faces of many immunologists. Clearly not a lady to be thrown off course by such reactions or by what, to many, would surely be an unnerving number of grant refusals, her research continued in the same vein throughout the period covered in the book.

Much of what is recounted rings true; for example, the long periods Anna Brito spends reading around her subject hoping to place her observations into a broader context. Other features of her laboratory life seem less convincingly told. Only twice, for example, is there any mention of her having any scientific competitors. This is either true and untypical, or untrue and misleading, given the extent to which competition drives so many scientists.

The chief virtue of the book is that it recounts in detail the thoughts and actions of a scientist brave enough to be open about her mistakes and failures and to bare many of her thoughts, despite the risk that some might look naive at the time and worse in retrospect. But the book fails to shed much light on the process of scientific discovery or to sustain interest for one simple reason; as it happens, nothing that I could classify as a substantial discovery emerges during the period covered in the

book.

Finally, potential readers might be attracted by two puzzles. The first, not too taxing, is the real identity of Anna Brito. The second, perplexing, is why, given the circumstances of the initial discovery that set her on the pilgrimage of research, Brito

can later say: "I can't stand people who just put everything together and try to see what happens — who don't think before doing an experiment . . . I could never do science like that"?

Peter Newmark is Deputy Editor of Nature.

Will the real Beaker Folk please stand up?

Colin Burgess

The Beaker Folk: Copper Age Archaeology in Western Europe. By Richard J. Harrison. Pp.176. ISBN 0-500-02098-1. (Thames & Hudson: 1981.) £12, \$19.95.

"BEAKER Folk", like the Wee Folk, exist in the mind of the beholder, and Dr Harrison is as uncertain in his attitude to "Beaker Folk" as most of us are about fairies. His last chapter, "Was there a 'Beaker Folk'?", was eminently sensible, and his conclusion that they "have no substance as a special population group", and that the Bell Beaker phenomenon began "as a fashion in fine pottery", would meet with widespread approval among students of the subject.



Gorge of the River Verdon in Provence, "an exceptionally interesting site".

Unfortunately, this attitude does not square with the contradictions in his preceding six chapters. The author claims in several places not to believe in "Beaker Folk", but clearly finds it difficult to erase them from his subconscious. His inner confusion will leave the reader feeling profoundly uncomfortable, because it is obtrusive throughout: to quote from the jacket, "the so-called 'Beaker Folk' were not a separate, well-defined community, but constituted more of an influence or movement", yet in the next sentence "They appear to have been instrumental in introducing metal technology", and their arrival "coincided with major changes in the structure of primitive society". Such ambivalence left me wondering at the end whether the author had experienced a late

change of heart about "Beaker Folk", and been forced to take a scalpel to his text; or perhaps it is simply that he, like most of us, finds it difficult to adjust to a world without "Beaker Folk".

In many ways, Harrison engineers his own downfall. On p.13 he complains that "One bad habit which many archaeologists have had is to abstract beakers from their local settings and study them as something alien and isolated", which is exactly what has bedevilled study of this period in Europe, and is exactly the mistake this book perpetuates. For sadly, Harrison is largely concerned with his title, *The Beaker Folk*, and hardly at all with his subtitle, *Copper Age Archaeology in Europe*. But until the former is strictly subordinated to the latter, understanding of this exciting phase of European prehistory will advance very little.

There is a fundamental controversy about Beakers at the moment, and it should therefore be a priority in a work such as this to begin with a fair statement of the issues. As it is, for much of the book — until the last chapter in fact — Harrison skates over the difficult question of Beaker interpretations with brief, throw-away allusions, and in so doing is hardly fair to those who have been trying to open up the whole question to review. For example, on p.11 he is decidedly disparaging about the suggestion that Beakers may have been central to a beer drinking cult, but on p.15 they may "have been status symbols . . . perhaps on account of a special drink", and on p.69 "Bell Beakers could also have derived their importance from their contents, such as beer, mead . . .". If he looks again at what S.J. Shennan and I wrote in *British Archaeological Reports* in 1976, he will see that our concern was not to set up alternative models to explain Beakers, but to show that there are explanations, other than ethnic ones, which will fit the evidence. Harrison, like some others, has seized on the evocative peyote/beer suggestion, but does not mention the equally relevant Butt Beaker analogy.

The author's preoccupation with the Beaker aspects of this phase of prehistory results in much of the book being a rather dreary regional catalogue of contexts and sites in which Beakers occur. How much more illuminating it would have been had he remembered his sub-title, and con-

sidered the local background. As it is the Beakers too often float in a vacuum. The reader will find it nearly impossible to grasp how they relate to regional cultural developments, for frequent allusions make it clear that Beakers were absorbed by local populations and adapted to suit local circumstances. Thus Myrthøj in Jutland is a "Bell Beaker settlement" (p.35), but only one house had Beakers. Local Late Neolithic pottery dominates, but where does it fit in? In Hungary, the Beaker is intrusive in local Vučedol contexts, while over the Eastern Beaker province as a whole "between 80% and 90% of all pottery found in Bell Beaker graves [sic] is not actually of a beaker shape at all [my italics], but small cups, jugs, bowls and jars". So why should these be "Bell Beaker graves" and Myrthøj a "Bell Beaker settlement"?

What is needed here, if ever we are to make sense of the Copper Age in Europe, is a radical change in thinking. These are not "Bell Beaker" graves and settlements, but the graves and settlements of whatever local cultural group is involved, which happen to incorporate Beakers and Beaker influence to varying degrees. Usually Harrison at least acknowledges a local cultural element, for example in southern France (pp.114-116), and especially in Italy (p.123), where recent discoveries offer a good chance of putting Beakers into proper perspective. Alas, in Britain and Ireland, Harrison opts for the traditional Beaker invaders. "Bell Beaker settlements", as at Newgrange and Knowth, are mentioned without making it clear that only a fraction of their pottery is classic Beaker; a lot of it is decidedly *sui generis*, and the rest local in character. For me, the treatment of Britain is particularly disappointing: the author fails to examine objections to a distinct Beaker physical type; he does not scrutinize critically the so-called Beaker innovations; and he appears not to appreciate the supreme importance of the Lanting and Van der Waals regional approach, not just for British Beakers, but for later British prehistory generally.

It would be churlish not to compliment the author on his choice of illustrations, which make clear the skill and patience which went into producing Beakers; and as a panoramic view of the variety of regional Beaker styles this book will be invaluable. But Harrison's aim is to "give an intelligible account of the Copper Age in Western Europe" (p.6) and he does not seriously attempt the task. It is a perilous time to write on the Beaker phenomenon, but suicidal not to have a clear idea in one's own mind beforehand of what Beakers are all about. I confess that I would not be so rash as entirely to dismiss the Wee Folk; but I have no such qualms about the "Beaker Folk". □

Colin Burgess is Staff Tutor in Archaeology, Department of Adult Education, University of Newcastle upon Tyne, and author of *The Age of Stonehenge* (Dent, 1980).

Mummies: new light on ancient societies

A. Rosalie David

An X-Ray Atlas of the Royal Mummies. Edited by James E. Harris and Edward F. Wente. Pp.403. ISBN 0-2263-1745-5. (Chicago University Press: 1980.) \$55, £33.
Mummies, Disease and Ancient Cultures. Edited by Aidan and Eve Cockburn. Pp.360. ISBN 0-521-23020-9. (Cambridge University Press: 1980.) £25, \$45.95.

GRAFTON Elliot Smith, Armand Ruffer, and Alfred Lucas led the field in the application of their disciplines — anatomy, histology and chemistry — to the study of ancient Egyptian mummified remains in the early part of this century. Since then, there has been a continuing interest in the use of medical and scientific techniques to amplify historical and literary knowledge of family relationships, the existence of disease, living conditions and funerary practices in this ancient civilization. Recent technological advances have encouraged further multi-disciplinary research, and the study of mummies has provided the Egyptologist with new information.

The discovery of the Egyptian royal mummies of the New Kingdom period (c. 1675-950 BC) was an archaeological find of considerable significance. The pharaohs and their queens, buried with great ceremony in rock-cut tombs at Thebes in Upper Egypt, were soon disturbed by robbers; their tombs were plundered and their mummies stripped of precious amulets. The high-priests of the god Amon re-buried many of these royal mummies in two resting places — the shaft of a tomb near Deir el-Bahri and a side chamber in a King's tomb in the Valley of the Kings — in the vicinity of Thebes. Here, they remained undisturbed until the chance discovery of the two caches in AD 1871 and AD 1898.

The mummies were subsequently removed to the Cairo Museum, where they were unwrapped and investigated by archaeologists and anatomists in 1889 and in 1912, when G. Elliot Smith produced his definitive anatomical work entitled *The*

Royal Mummies. This included the examination of kings, queens, princes, princesses, priests and some unidentified individuals. However, having the opportunity to examine only one of these mummies with radiographic techniques, he stated that "... examination with the aid of X-rays would, no doubt, have provided much additional information and I hope this will be done at some future time ...".

This opportunity did not arise until 1967, when the University of Michigan, which had previously collaborated with Alexandria University in a radiographic survey of the skulls of an ancient Nubian people, was invited by the Egyptian authorities to X-ray the royal mummies in the Cairo Museum.

Their aims were to add to historical and literary sources relating to the chronology of this period and to the lengths of the rulers' reigns; to assess the health of each individual; to check the reality of the royal sculpture and art forms against the physical evidence of the mummies; and, using a portable X-ray cephalometer, to determine the craniofacial variations and the degree of physical heterogeneity evident in the pharaohs and queens of the New Kingdom.

The *Atlas* provides the long-awaited reference work (a popular interim account by J. E. Harris and K. R. Weeks, *X-Raying the Pharaohs*, was published in 1973 by Charles Scribner's Sons). It does not give a detailed description of each mummy, but, in addition to the illustrative material, uses a series of self-contained chapters to highlight features of this unique collection. The chapters reflect the individual interests of the multi-disciplinary team, and include techniques of mummification, the practice of medicine and dentistry in ancient Egypt, evidence of physical and dental disease in the royal mummies, radiological techniques and findings, the genealogy of the Royal Family and craniofacial variation, and the age at death of the New Kingdom pharaohs (here, Egyptological and



Mummy of King Ramesses II, nineteenth dynasty.

From X-Ray Atlas of the Royal Mummies

physical evidence, sometimes at variance, is assessed separately). Such an approach, as admitted by the editors, results in a degree of repetition and an overlap of material in some instances. However, it is useful and interesting to see the physical and historical evidence summarized independently.

Another difficulty which the investigators faced, because of the paramount consideration for the safety of the royal mummies, was the necessity of examining the mummies in their modern wooden cases in the museum, instead of under near-ideal conditions in a modern radiology department where sophisticated equipment and a variety of radiological techniques would have been available. There is a consequent lack of quality in the X-rays and possibly in the completeness of the evidence obtained.

However, some of the most interesting results included the observation that, unlike certain other Egyptian mummy collections, these royal mummies showed a marked degree of cultural and genetic heterogeneity. This reflected the presence of foreign princes and princesses in the Egyptian royal family of the New Kingdom rather than the averred custom of royal consanguineous marriages undertaken to ensure succession to the throne. Another tentative conclusion, reached through the use of scientific techniques, was that a mummy known previously as the "Elder Lady" was in fact probably Queen Tiye, a close relative of Tutankhamun.

Nevertheless, the caution which historians should observe in utilizing the results of such scientific investigations is well illustrated by one of the conclusions drawn here. This concerns the proposed relationship between the pharaohs Amenhotep III and Tutankhamun. Previous investigations of the mummies attributed to these kings, undertaken by Professor R. G. Harrison and his colleagues, and based on palaeoserological and anatomical observations, had indicated that Amenhotep III was probably Tutankhamun's father. However, a different conclusion is drawn here: "The cephalometric X-ray records would not support the . . . contention. It is most improbable that the mummy designated Amenhotep III was the biologic father of King Tutankhamun".

The second book reviewed here, *Mummies, Disease and Ancient Cultures*, deals with a wider range of techniques and also shows how a variety of societies have preserved their dead, either unintentionally or by artificial methods, as "mummies". This study again shows the value of a multi-disciplinary approach. In the long term, it is hoped that the work will go some way towards answering questions concerned with the origin of diseases and their evolution in different societies. To this end, the team is committed to a detailed pathological study of world-wide collections of mummies.

A summary is given of the work to date, and although Egyptian and South American mummies have received most detailed attention at present, a survey of other mummified remains is included — those discovered in the United States, the Aleutian Islands, Alaska, the bog bodies of Denmark, mummies in Australia and Melanesia, and Japanese examples.

A complete section is devoted to reports of the techniques involved in this study which include physical anthropology, radiology, histology, electron microscopy, palaeoserology and palaeobiochemistry. Not only does the book illustrate the various methods — heat-drying, freezing or the application of dehydrating agents —

by which a body can be preserved, but it also shows how diverse are the religious and environmental reasons which bring about mummification in different parts of the world.

Both books reflect a resurgent interest in the application of scientific techniques to the physical remains of ancient societies, which, used with due caution and appreciation of the possible fallability and limitations of such investigations, can provide the historian with a valuable additional insight. □

Rosalie David is the Egyptologist at the Manchester Museum, University of Manchester, and Director of the Manchester Egyptian Mummy Research Project.

The cradle of geology refurbished

John Sutton

The Geology of Europe. By Derek V. Ager. Pp.560. ISBN 0-7084-115-2. (McGraw-Hill: 1980.) £16, \$38.40.

IN MY opinion this is much the best account of the geology of Europe to have appeared since Gignoux published his *Géologie Stratigraphique* over 50 years ago. Derek Ager describes his subject as Europe seen by a stratigrapher-palaeontologist. He treats his material very differently from Gignoux and the contrast is illuminating. Gignoux drew nine-tenths of his examples from Europe, and, in so doing, produced a comprehensive account of its geology which started with the oldest rocks, proceeded through younger formations to the most recent, providing abundant stratigraphical successions and the like on the way. Ager throws this approach out of the window; there is hardly a stratigraphical table in the book, so that he has space for much else.

Ager divides Europe into four regions on the basis of the relative ages of the last extensive period of metamorphism and deformation to have affected them. The geological history of each region is then described from start to finish — a method which brings out relations of older and younger events particularly well. Each of the four sections ends with a concluding chapter a few pages long. I was particularly struck with Ager's concise, five-page summary of Precambrian Europe. Although, as he puts it, "the Precambrian history of our continent is for the most part lost in the mists of metamorphism", he packs a telling account into that brief chapter.

The main virtue of this book is the way in which it brings together details to give a geological picture on a continental scale. Derek Ager has an interesting approach to geophysics; he appreciates "the hurricane of fresh air that new geophysical thoughts breathed into our subject". He goes on to write what is one of the best accounts of

any continent, set out on precisely the scale which would make it possible to compare the geological history of basins and uplifted crystalline blocks with the results of geophysical surveys. Such comparison, however, he leaves to others. He gives a good defence of his treatment, and in his introduction makes the point that any additional matter has to be balanced by the omission of something else. However, in a second edition I would hope at least to see a discussion of outline gravity and magnetic maps, and an account of crustal thickness.

Of all the continents, Europe is probably the most difficult to handle. Its geology is discussed in 20 languages, it is criss-crossed by frontiers and its complexities are probably the most apparent of any continent. Ager sails through all these difficulties triumphantly. He has travelled the continent repeatedly, he lists friends and helpers in every country; he is bold in his summaries and he writes with great clarity, and is not afraid to say what he thinks is important. Personal touches abound; we hear about the view from his study window and the nature of the rock surmounting the Ager garden gate. Historical figures pop up from time to time: Drake's game of bowls and the sailing of the Pilgrim Fathers illuminate the account of Plymouth Hoe. Whether, in consequence, one remembers which stage of the Devonian outcrops below the turf is probably a matter of individual experience. I expect every reader will find some of the allusions obscure; there are plenty of traps for the translators who will certainly get to grips with this most useful book. Unless they know their Sherlock Holmes and Irish mythology, we may yet see *chien de chasse des Basquervilles* and *Blarneystein* among the faunal lists and formation names of the future. □

John Sutton is Pro Rector and Professor of Geology at Imperial College, University of London.

Clocks in grandfather's rocks

G.Y. Craig

The Abyss of Time: Changing Conceptions of the Earth's Antiquity after the Sixteenth Century. By Claude C. Albritton Jr. Pp.251. ISBN 0-87735-341-7. (Freeman, Cooper: 1980.) \$12.75 US only.

THE essence of this story is quite dramatic. The calculated age of the Earth has increased from a little less than 6,000 years to more than 4,500 million years in the last three centuries. The Earth, of course, has remained untouched by this spectacular revolution apart from the addition of these few hundred years. What *has* changed has been the consensus of opinion by philosophers, theologians, physicists, geologists and evolutionists as to how old the Earth actually is; and it is salutary to be reminded by younger colleagues that estimates of the age of the Earth have doubled even since I was a student in the 1940s. Most fortunately this precipitous rate of ageing seems to have stopped.

The two main ways of looking at the passage of time in geology are not dissimilar from the ways in which we recall our own lives. Events can be put in a rough sequence and then, often with difficulty, slotted into the calendar ("Remember that

The discovery of radioactivity at the end of the last century led to increasingly precise radiometric dating of those rocks containing suitable isotopes; and the two scales, relative and radiometric, can now be correlated to provide a chronological scale of historical geology. This scale is still being extended and refined. Metamorphic rocks in Greenland, for example, have recently been dated at 3,750 million years, the oldest rocks so far discovered.

Dr Albritton's account arose out of a series of lectures given to students and staff at various academic institutions in the United States. The slim, attractively bound text follows the traditional and by now well-worn path through the maze of geological time by examining the contributions of its protagonists. Leavened by portraits and anecdotes, it is easily read and clearly reveals the influence of religion, conceit, catastrophe, fanaticism, feuds, ignorance, intuition, rocks and radioactivity on the evolution of the concept of geological time.

In spite of jazzy chapter headings — Malta's cradle; Solids within solids; Mr Hook(e); A sacred theory of the earth; Telliamed's story; The epochs of nature; View from the brink; Father William; The four-dimensional jigsaw puzzle; A question of tempo; A plenitude of events; The great tree of life; Kelvin; A numbers game; Radiometric dating; World enough and time — the leading actors are readily unmasked. For the reader who does not care to play this game, the protagonists turn out to be Steno; Hooke; Burnet; De Maillet; Hutton; Smith; Werner and other early stratigraphers such as Sedgwick and Murchison; catastrophists and uniformitarians (I've now broken a New Year resolution not to use that word!) especially Cuvier and Lyell; observers of volcanoes and glaciers (Scrope and Agassiz); Darwin; and lastly late-nineteenth and twentieth century scientists. Pleasingly, the book includes an appraisal of Arthur Holmes' brilliant contributions to the construction of a geological time-scale. The last chapter contains a typical student-textbook analogy: by compressing the major events of the Earth into 365 days, primates appear on Christmas Day and *Homo sapiens* arrives a mere hour before the dawn of the New Year. Dr Albritton is obliged to devote part of this chapter to a rebuttal of creationist dogma, now so common in parts of the States. Such concern is understandable for he teaches in Texas where fundamentalist views are not unknown.

Yet somehow the excitement of what should be a gripping story is not there. In one form or another it has all been done before, albeit with the historical anecdotes cut to a minimum. Last year two long articles by John MacPhee in, of all places, *The New Yorker*, conveyed a similar



Mary Evans

William Thomson, Baron Kelvin (1824–1907). Detail from a cartoon in *Vanity Fair*.

summer when we went to Brittany and Aunt Doris got . . . Now what year was that?). And so it is with rocks. Geologists during the nineteenth century pieced together a sequence of geological events which enabled them to construct a relative time scale. Concurrently various attempts were made to establish when these events occurred and especially how old the Earth might be. Such calculations involved rates of erosion and deposition, the rate at which salt is being added to an already salt sea and the rate at which the Earth appeared to be cooling.

SCIENTIFIC BOOKSHOP

H.K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.50. (Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

H.K. LEWIS & Co. Ltd.

136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavit,
London, WC1E 6BS."

Circle No.19 on Reader Enquiry Card.

Bring Your Library Up-to-Date With This ANS Edition ... Nuclear Energy in Germany

Winnacker/Wirtz
6 x 9 hardbound edition
\$37.00

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Please take my order for _____ books.

- ☐ Payment in full enclosed.
☐ Bill me and I accept postage and handling charges.

Name _____

Org. _____

Street _____

City _____

State _____ Zip _____

Circle No.02 on Reader Enquiry Card.

message but in a more entertaining style, perhaps because his serious but flowing story was laced with irrelevant cartoons and adverts including fast cars, whisky, cameras, jewellery and beautiful women, but more probably because he linked history with current field-work in different parts of the States.



James Ussher, Bishop of Armagh — calculated the date of creation as 4004 BC.

So here I am, vaguely dissatisfied with a perfectly adequate book simply because of a feeling of *déjà vu*. How can one be constructive? More might have been made of Archbishop Ussher (1581–1656) and his method of calculating the date of creation (4004 BC) and such subsequent events as the grounding of the Ark on Mount Ararat on May 5, 1491 BC (a Wednesday!). His chronology was incorporated in the most widely read book of the day — the King James VI bible — and so had an undue influence in Protestant countries. Many of the subsequent philosophical and scientific investigations described in the *Abyss of Time* were carried out in the shadow of Ussher's precise chronology. Astronomical calculations and coral clocks are not referred to but they show that the Earth is now rotating more slowly than it did during the Devonian when a year was all of 399 days. Above all, data from meteorites and Moon rocks reveal fascinating events earlier than any now preserved on Earth and provide independent factual evidence of the great antiquity of our Solar System and planet. They could have been the subject of a separate chapter.

And what of the future? The abyss is just as much in front of us as behind. The Earth is steadily losing its heat. As it continues to cool the lithospheric plates will thicken and stop moving: eventually there will be no more volcanic activity and no more earthquakes, but the land will continue to be eroded seawards, there to remain because uplift will no longer occur. The rock cycle will finally have come to an end: but there is perhaps consolation in the fact that the Sun will continue to shine until in another 5,000 million years it too becomes cold. And after that? But Albritton does not indulge in such fantasies. □

G.Y. Craig is Professor of Geology at the University of Edinburgh.

Trap for the unwary

William H. McNeill

The Fates of Nations: A Biological Theory of History. By Paul Colinvaux. Pp.383. ISBN 0-671-25204-6. (Simon and Schuster: 1980.) \$12.95 US only.

FOR the intellectual wanderer, impatient with the myopia of specialization, there are great risks in invading an unfamiliar turf. Jealous experts lie in wait for the newcomer to display his ignorance of their discipline by some slip of tongue or pen, and the complications of the new subject may not conform readily to the insights the newcomer brings from some other discipline.

This, unfortunately, is the case with Paul Colinvaux's book, *The Fates of Nations*. Colinvaux is an ecologist whose speciality is Ice Age Alaska and the land bridge which once allowed animals and plants to migrate from Asia to America across what is now the Bering Strait. For general readers, however, his best-known work is a summation of recent ecological niche theory which goes by the catchy title *Why Big Fierce Animals are Rare* (recently issued in paperback by Penguin). As an historian, who, like Colinvaux, has found it irresistible to wander across recognized disciplinary boundaries, I found this essay instructive and persuasive. But now when he takes on my own field of expertise, the human past, and seeks to illumine it by bringing ecological niche theory to bear, I find Colinvaux's ideas inadequate and his information radically defective.

As he engagingly explains at the end of the book in a "Note on Sources and Substance", he began to think about a "formal theory of history" in 1971, relying on the history "which is the common knowledge of Christendom", i.e. on what he remembered from his school days in England. This he reinforced by reading Toynbee's "overview of history's facts", and by consulting a few experts on battles — Livy, Tacitus, Arrian — as well as more modern writers like Liddell Hart and J. F. C. Fuller. He found the barrister Edward Creasy's *Fifteen Decisive Battles of the World*, first published in 1854, especially useful, while Field Marshal Montgomery's *A History of Warfare* supplied him with "the clearest insight into what gave victory in many a crucial battle". Thus, as Colinvaux tells us himself, his method of mastering history was studiously to disregard what professional historians have been doing since World War I to modify the patriotic and political history so ably caricatured in *1066 and All That*.

With *1066 and All That* plus school-boy Latin as his guide, and Toynbee as guarantor of "facts", what does an intrepid ecologist discover in the human past? The answer is a rhythm of rise and fall, predictable in the light of ecological niche theory. It may be summarized as

follows. Initially, history starts when a group of human beings develops a new technique that allows them to enlarge the ecological niches in which they live. This leads to a rise in population, for human beings have always pursued a breeding strategy — the large young gambit — that requires each breeding couple to raise as many children as they think they can afford. This strategy, Colinvaux assures us, has never altered, though he does not assert that it is instinctive.

Once enlarged niche spaces start to fill up as population increases, the prospering community faces a crisis. Since growing numbers with no increase in resources will force them to accept narrower niches again, they must look abroad to find the means to sustain their heightened expectations. Trade, colonization and aggressive war, aimed at taking land away from those in possession, are the alternatives such a population faces, and this is what history is all about. Successful communities expand and create empires; but sooner or later, expansion ceases. Colinvaux is completely silent on what stops expansion. On his principles I can see no reason for the process he sketches not to continue indefinitely, until the most successful aggressor has created an empire of the Earth. But the "common knowledge of Christendom" makes clear that a limit was set somehow, and Colinvaux shows his deference to facts by stopping the expansion of the Roman Empire at the Rhine and Danube — though, as I say, he gives no reason for this interruption of a process which he has declared to be both predictable and inevitable.

Once territorial expansion stops, the continued growth of population requires narrowed niches; and narrowed niches will be accepted only reluctantly by the population concerned. This means an oppressive government — a police state in fact, in which the rich, clinging to their broad niches, impose their power on the poor by dint of "regimentation, bureaucracy, class, rationing and caste" (p. 64). Such a regime is vulnerable from without, and eventually barbarian invaders, technologically abreast of the latest methods of civilized warfare, will succeed in destroying the oppressive regime by cutting off food supplies to the cities. The newcomers institute a barbarian pattern of life in which local consumers depend on local food production. This sharply reduces the total number of inhabitants, since the urban masses, congregated under the regime of civilization into vast cities, are unable to maintain the breeding strategy that had hitherto led to an ever-mounting population. They simply die off, leaving a sparsely settled countryside to start the cycle of civilization over again.

Echoes of Toynbee's cycle are obvious, though the emotive tone that Colinvaux's ecological vision imposes is utterly at odds with Toynbee's humanistic and religious values. There are, too, some serious anomalies in this thesis, for the modern cycle of civilization does not conform closely to the classical Mediterranean pattern.

An even greater anomaly is introduced in Colinvaux's final chapter "The Shape of Things to Come". After exploring the implications of this theory, which predicts that growing populations will launch aggressive war in order to maintain enlarged ecological niches, he apparently reverses himself by saying that it "may well be that our breeding strategy will yield even smaller families for the moderately affluent than they have now" (p. 313), since one may hope that "new roles for women will hold down the population and let us all live in freedom" (p. 314). So because "we breed in human and controllable ways" and can understand what is happening to us, we "need fear neither our future nor our fate" (p. 321). But if human breeding strategy is alterable now, why not in the past? And if one group of people in this twentieth century does decide to have fewer children in order to maintain broader niches, what prevents others from breeding incontinently and filling in the space occupied by the rich and continent, not by conquest, but by the less spectacular method of in-migration?

This sort of population shift is occurring before our eyes, as *Gastarbeiter* in Germany, illegal aliens in the USA and Asians in England attest. It occurred in distant ages as well. The replacement of Sumerian speakers by Semites in ancient Mesopotamia in the 3rd millennium BC and the bitter jibes of the Roman poet Juvenal against the Greeklings whom he saw taking over favoured places in Roman society, are only two examples of the past importance of peaceful migration. Indeed, all great imperial cities of times past have been polyethnic in their population. They were places where diverse peoples congregated and died off due to the intensified disease exposure involved in urban living before the days of modern sanitation. Colinvaux seems unaware of this traditional population dynamic, well attested though it is for early modern Europe when statistical records begin to become available to historians. Indeed he is ignorant of all historical demography, though for a theory that puts such emphasis on population growth as the fundamental disturber of history he surely ought to have guarded himself against the sort of reproach I make of his procedures by at least consulting a few elementary textbooks on historical demography. If he had done so, he would have discovered that historians are agreed that the population of the Roman Empire declined sharply in the second and third centuries AD, recovered partially in the fourth century and then plummeted again in the fifth. Colinvaux's

picture of an ever-rising population pressing harder and harder on available niches therefore runs counter to all modern historical scholarship.

As a matter of fact, Colinvaux is aware of some of the signs of depopulation that manifested themselves in late Roman times. But he dismisses empty fields and unoccupied border regions where Roman authorities settled barbarian tribesmen from beyond the frontier as a predictable consequence of population growth! The argument is sufficiently bizarre to deserve quotation in full:

From the time of troubles onward the population shifts within the state. There is a drift from the land as peasants are displaced in the interests of increased production. This happened in ancient Greece and Rome no less than in the time of the enclosures in Tudor England, or in modern industrial states. The process can be seen in the development of every civilization. Feeding great numbers of people is more easily done if they are brought together in dense settlements; running the agriculture needed to supply these dense settlements is more efficient in the larger agricultural units of agribusiness. And the "drift from the land" is a predictable consequence of a dense population growing denser [pp. 83-84].

To test this hypothesis one might look for some indication of the growth of cities while fields emptied out. Roman cities from the 3rd century onward conveniently provided such an index by erecting ring walls necessitated by the declining level of public security in the empire. By modern standards these enclosures were tiny indeed, and diminish with each rebuilding. Where is the massive, ever-growing urban population Colinvaux's theory requires? Even on a purely theoretical level, and without the vulgar appeal to mere data, one would suppose that before propounding his ridiculous hypothesis Colinvaux might have asked himself how it could be more efficient to carry food from the fields to distant urban consumers (in ox carts presumably), and how "agribusiness" could plough and reap in the absence of modern machinery, and without numerous labourers on the spot to do the work. The fact is that human populations rise and fall in complicated and only partially understood ways, and in response to a far wider range of shifting conditions than Colinvaux's system admits.

My judgment of *The Fate of Nations* is that in venturing beyond the boundaries of his discipline Colinvaux has come a cropper. This is a pity, for surely there are useful things historians might learn from ecologists, and an ecologist who took the pains to acquaint himself with historical literature might indeed be in a position to shed new light on familiar historical conundrums.

William H. McNeill is Professor of History at the University of Chicago, and author of *The Human Condition* (Princeton University Press, 1980).

THE EARTH'S PROBLEM CLIMATES

Second Edition
by Glenn T Trewartha

In this new and extensively revised edition, the author takes into account new information and climactic data that have become available since the original edition appeared in 1961.

Trewartha offers an inventory of the more important of the earth's problem climates, brief and concise descriptions of their characteristics, and suggested explanations for at least some of them; in so doing, he draws together the results of numerous definitive studies in dynamic climatology.

Wisconsin, May 1981, 352 pp, illus, cloth £13.50

VOLCANOES OF THE EARTH

by Fred M Bullard

"... a first-rate and up-to-date account of one of the world's most intriguing phenomena."

-*Explorers Journal*

Texas, 597 pp, 24 colour plates, 183 b&w illus, now in paperback at £12.00

AMERICAN UNIVERSITY
PUBLISHERS GROUP LTD

1 Gower Street
London WC1
Telephone (01) 580 3994

Circle No.41 on Reader Enquiry Card.

SCIENCE AT THE WHITE HOUSE

A Political Liability

EDWARD J. BURGER, JR.
foreword by DON K. PRICE

From the day the first atomic bomb was dropped, scientists gained a new image in the public's eye. In the succeeding decades, there seemed to be no limit to what technology could accomplish, from military weapons to moon walks. *Science at the White House* looks at what happened when America's new experts were called in to serve as advisers to the president.

As cases in point, Burger assesses the government's efforts at planning for key issues that emerged during his tenure as a science adviser: issues concerning national health care, environmental regulation, and population and family planning, among others. 208 pp, £9.00.

THE JOHNS HOPKINS
UNIVERSITY PRESS

Ely House, 37 Dover Street,
London W.1.

Circle No.47 on Reader Enquiry Card.

One hundred years of the Natural History Museum

D.T. Donovan

Alfred Waterhouse and the Natural History Museum. By Mark Girouard. Pp. 64. ISBN 0-300-02578-5. (Yale University Press: 1981.) Hbk \$12.95, £4.95.



Alfred Waterhouse, 1830–1905, from a portrait by Sir William Orchardson.

THE Natural History Museum is the most splendid of the South Kensington museums; it has a vigour and coherence which is a tribute to the architect Alfred Waterhouse. So it comes as a surprise to learn that the building has a complicated history, and was nominally designed by somebody else.

The natural history collections were for long kept in Bloomsbury with the books and antiquities of the national museum. By the early nineteenth century antiquities were being added in large quantities, and by mid-century many thousands of natural history specimens were accumulating every year. The new British Museum building, begun by Robert Smirke in about 1823, was already too small by the time it was completed in about 1850.

Mark Girouard begins his account with

Richard Owen's vision of a great new museum in 1858, which was the origin of the present building. Removal of the natural history departments from Bloomsbury had been considered for some time, though a practical solution was not found until after 1851 with the development of South Kensington as a centre for education in the arts and sciences.

Owen, already famous for his contributions to comparative anatomy, had been appointed Superintendent of the Natural History Departments in 1856. In the same year, Antonio Panizzi, Keeper of Printed Books, was made Principal Librarian (i.e. Director) of the museum. Panizzi had long wanted to get rid of the natural history departments, and said so on every possible occasion. Yet even with Owen and Panizzi working towards the same end, it was another 25 years before the natural history collections were actually moved out of the Bloomsbury building. Initially, at least, almost the whole of the scientific establishment was opposed to such a move, and succeeded in influencing Parliament against it for some time.

In 1863, however, the Government bought the present site. On it stood the exhibition building of 1862, built by the official architect Francis Fowke, and widely disliked. Parliament refused to purchase this building, which was demolished, and an architectural competition for a great new museum complex was held in 1864. It was won, to everyone's surprise, by Francis Fowke, but he died before his plan could be executed.

Alfred Waterhouse, newly arrived in London from successes in Manchester, was appointed to carry out Fowke's design, but it was not until 1873 that building was started. By this time only the Natural History Museum — about one-quarter of Fowke's complex — was actually going to be built, so the building had to be drastically re-designed. Girouard traces the

influence of Fowke's design by comparing successive versions of the building. However, there is little perceptible resemblance between Fowke's drawings and the present building, and Waterhouse's standing as the architect is fully justified.

The plan of the museum, however, is neither Fowke's nor Waterhouse's; it stems from a sketch by Owen in 1859. This shows main east–west galleries with numerous galleries opening off them at right angles, the arrangement in the present-day museum.

Owen had always wanted his museum to be enriched by appropriate decoration. He and Waterhouse devised a scheme in which representations of living animals and plants adorned the west wing, intended for zoological collections, and extinct ones (restorations, not fossils) the east wing, where fossils and minerals were to be housed. Waterhouse's drawings were translated into models and cast in terracotta; comparative pictures in the book show how faithfully this was done.

The book has been sponsored by the Museum as part of the celebration of the centenary of its opening in 1881. This centenary is, unhappily, marked by controversy over the current proposal to demolish the rear single-storey galleries, thus destroying Owen's original plan.

Mark Girouard has written an attractive and interesting book which explains the history of the building and its architectural origins. It is well illustrated with pictures of the museum as it is, Waterhouse's working drawings and earlier versions of the design. Only the photographs of the interior are disappointing and do less than justice to the building. Otherwise, the book makes a handsome souvenir of this great embodiment of Victorian scientific achievement. □

D.T. Donovan is Professor and Head of the Department of Geology at University College, University of London.



As it might have been — the winning, but unexecuted, design submitted by Francis Fowke in 1864. On the death of Fowke, Alfred Waterhouse was appointed as architect and the museum was built largely to his designs.

Intellectual cross-fertilization and evolution

A.J. Cain

The Evolutionary Synthesis: Perspectives on the Unification of Biology. Edited by Ernst Mayr and William B. Provine. Pp.487. ISBN 0-674-27225-0. (Harvard University Press: 1980.) \$25, £15.

THIS book looks out upon an incredibly rich panorama of personalities, groups, institutions, countries, attitudes, theories, arguments, experiments, facts, false facts (even a very few artefacts), misunderstandings, barriers emotional, linguistic, dogmatic and intellectual, all adding up to one of the greatest intellectual advances of our time — the unification of different biological sciences by a single neo-Darwinian theory of evolution in the first half of this century. Some of the principal actors have contributed invaluable accounts of their own roles: among them are Mayr, Simpson, Dobzhansky, E.B. Ford, C.D. Darlington, Curt Stern; and there are extremely valuable surveys and analyses by biologists, historians and philosophers (Lewontin, Gould, Mark Adams, Dudley Shapere etc.). The book itself is now history and, since the 1900–1940 period qualifies as history (not merely as out-of-date), it will inspire and provoke an ebullition of PhD theses, papers and books worse than the great Wittgenstein industry. One of the virtues of the book is that while making thoroughly clear the limitation of science by the inadequacies of personalities, institutions and governments, it does present beautifully, especially in genetics and cytology, the intellectual impact of actual scientific discoveries and their role in changing people's thinking.

The book is organized by subjects (Part I, cytology, embryology, systematics, botany and so on) and by countries (Part II, Russia, Germany, England, France, USA), with a concluding survey (no, I will not say "overview") of the whole subject and a miscellany of very useful biographical contributions, rather over-titled as "Biographical Essays".

Especially valuable to those with little German and less Russian is the survey by countries, and a saddening spectacle it presents; the superb beginning in Russia, based on an excellent naturalists' tradition of population genetics, aided at first by the ideals of education and especially science during the Revolution, then brutally terminated as a result of the charlatan Lysenko's power-politics. France did not contribute at all to the synthesis except for Teissier's and L'Heritier's invention of the population cage and work on balanced polymorphisms. Both started as mathematicians (the probable reason for Teissier getting his chair at the Sorbonne is fascinatingly non-scientific), uncontaminated by the ideas of the French Establishment.

Both Boesiger and Limoges have some very hard things to say about the academic establishment in France, and, so far as I can see, they are thoroughly justified. Boesiger remarks "France today (1974) is a kind of living fossil in the rejection of modern evolutionary theories; about 95 percent of all biologists and philosophers are more or less opposed to Darwinism."

Limoges comments equally astringently on the language barrier, the closed university system ("no one can hold a chair in a French university who is not French and who does not have a French doctorate") and the centralization of the system, providing for career control by a few senior professors in Paris. Although that has now been considerably modified, Boesiger emphasizes the centralized control of the research organizations. One can add that there is considerable pressure on French workers to publish in French. Certainly there are in France today some excellent and up-to-date evolutionists — one thinks immediately of Maxime Lamotte in Paris, and a botanical group at CNRS Montpellier, as well as Zuckerkandl and others at that university; nevertheless it seems to be true that a largely self-perpetuating Establishment, accepting only conformist successors, has made France, as far as evolutionary studies go, into an intellectual island. Yet the new recruitment every year of fresh minds is one of the most valuable processes in any intellectual society.

Germany is well treated by Mayr, Rensch and Hamburger. Mayr brings out important administrative limitations (not enough chairs and the nature of professional duties), Curt Stern the difficulties of institutions separated by distance (Berlin-Buch and Berlin-Dalhem!). Little is said about Nazi influence. Rensch, in an essay very sympathetic to his predecessors and colleagues, brings out many reasons for the slowness of the synthesis in Germany — apparently contradictory phenomena in genetics and palaeontology, speciation apparently concerned only with trivial characters, the apparent internal drive of the organism in development and regulation, independent of the environment, religious convictions and so on, which, however, operated elsewhere as well. The slowness of penetration of genetical results in Germany certainly needs study.

Hamburger brings out a further point (which Boesiger applies also to Cuénot in France): "the widespread German tendency to have a unified *Weltanschauung*, a more general overview of metaphysical and scientific creeds than just having *Weltanschauung* and scientific work compartmentalized side by side".

The history of the evolutionary synthesis in the USA is particularly well treated, and deservedly so, and there are many studies of that intriguing man T.H. Morgan and

Minerals from the Marine Environment

Sir Peter Kent, FRS

Resource and Environmental Sciences Series

A review of the full spectrum of mineral resources development as seen against the environmental background of the seas and oceans.

£3.95 paper 96 pages

The Ecology of Natural Resources

Second Edition

I.G. Simmons

The author has strengthened the new edition both by updating it statistically and by changing it conceptually. Increased emphasis is given to energy as a resource and as a link with other resource processes.

£8.50 paper 448 pages

Additions to the *Studies in Biology* series:

124 Light and Plant Life

J.M. Whatley and F.R. Whatley, FRS

£2.75 paper 96 pages

125 Lipids and Polysaccharides in Biology

Anna J. Furth

£2.10 paper 72 pages

126 Bioenergetics of Autotrophs and Heterotrophs

John W. Anderson

£1.95 paper 64 pages

127 Invertebrate Respiration

Rufus M.G. Wells

£2.25 paper 76 pages

128 Immunobiology

Christopher J. Inchley

£2.50 paper 80 pages

129 Growth Regulators in Crop Production

Leonard C. Luckwill

£1.95 paper 64 pages



Edward Arnold

41 Bedford Square,
London WC1B 3DQ

his close associates. The conversion of this school to a neo-Darwinian attitude by their own discoveries in genetics and karyology, powerfully aided by their openness to ideas and visitors from abroad, was, however, only partial in some cases because of the difficulties of extending genetical findings to macroevolutionary phenomena, including speciation. Mayr makes an especially important contribution in emphasizing that the synthesis was a synthesis; it was extremely difficult for many years to reconcile Darwinism and the new science of genetics, and it was systematists in particular who contributed the concepts of adaptive geographical variation, variation in whole populations under environmental selection, and partly adaptive speciation (a macroevolutionary process). Yet it must be added that it was largely the wholly unjustified insistence of many systematists that characters of species were non-adaptive that misled many of the earlier evolutionists into doubts about the efficacy of natural selection and the continuity of micro- and macroevolutionary phenomena.

Equally fascinating is the comparison between subjects. Botany (not the botanical part of genetics) was almost as poor a contributor among subjects as France among countries, and Ledyard Stebbins's article is as thought-provoking as Boesiger's or Limoges'.

When one stands back from it all, several major thoughts come to mind. The difficulties are highly instructive, but little more than one would expect in any intellectual controversy; indeed Simpson is probably right in indicating that given the human character, the synthesis was achieved about as fast as could be expected. But some topics are hardly touched on. I suspect (in a few instances, I know) that religious considerations came first, even if they were seldom expressed, in some scientists' minds, and conditioned heavily what they were prepared to accept scientifically. Under this heading I include humanism when it refuses the application of the synthesis to mankind, and of course Marxism with its leaning to some form of Lamarckism (only too well documented in this book).

Secondly, and allied to the preceding, the role of chance as an explanatory factor is hardly touched on — yet, on the one hand, as a means of explaining away inconvenient phenomena it has been invaluable as an alternative to natural selection; on the other, it has added to the revulsion of many people against natural selection that what the process has to work on is the effects of mere chance.

But lastly, there is almost no indication in the book that the recognition of natural selection as the major (or nearly exclusive) agent of evolution was ever due to any work on natural selection in the wild, except very indirectly. A simple-minded (or extremely intelligent) Martian might well think that if it was debated whether natural

selection operated in the wild, the thing to do was to go and look for it. Chetverikov might have done so if politicians had let him; Haldane thought it much too difficult. The whole debate over the efficiency of natural selection in the wild was conducted by argument over indirect evidence, and indeed to a large extent it still is.

The present book contains little material towards a history of actual field-work directly on natural selection, except for E.B. Ford's contribution — and perhaps rightly, since it seems not to have affected the synthesis. It does contain valuable hints towards a history of why no one has attempted it — for example, Lewontin's statement that "Wright's notion that random processes allow the exploration of the field of gene frequencies . . . frees the evolutionist from having to ask such

questions as why did some Ceratopsian dinosaurs have three horns and some have two and some have a frill and some not?". Of course it does nothing of the sort except to those who want it to. It was not *shown* in the first place that the development of different ways of fighting was due to any random process, except on the presumptuous assertion "Because I can see no sense in this phenomenon, not having worked on it, therefore it must be random". The use of random processes in explanation even at the present day is rather like John Henry Newman's use of St Augustine's text to precipitate him into the Roman Catholic church (see Geoffrey Faber's *Oxford Apostles*) — using a bad reason for doing what we intend to anyway. □

A.J. Cain is Derby Professor of Zoology in the University of Liverpool.

Avian evolution transcends philately

Alec Panchen

The Age of Birds. By Alan Feduccia. Pp.208. ISBN 0-674-00975-4. (Harvard University Press: 1980.) £12, \$20.

OF ALL those who concern themselves with the palaeontology and evolution of vertebrate animals, it is arguable that bird palaeontologists have taken on the most difficult task. The anatomical uniformity of the majority of bird species is such that the identification of their individual fossil bones varies between difficult and impossible; further, there are numbers of cases (some cited by Feduccia) in which the avian status of particular specimens is still a matter for debate. It is probably for these reasons that while popular books on dinosaurs are legion, a book on bird palaeontology is unusual despite the innumerable amateur ornithologists that are the mainstay of almost every natural history society. One would hope, therefore, that there was a market for a popular book like Feduccia's as evidence that at least some of these amateurs have interests extending beyond the prevalent stamp-collecting approach to bird watching.

On the whole Feduccia serves his potential audience well. However, his opening chapter tends to hop about its subject matter in a manner reminiscent of small birds themselves, starting with a scene-setting paragraph on *Archaeopteryx* — geologically the oldest as well as the most famous of all bird fossils — but then switching to bird anatomy, "Darwin's finches", adaptive radiation, the concepts of homology and analogy, the reconstruction of phylogeny and the geological time scale, before settling on an account of the several finds of *Archaeopteryx*, with useful photographs of each specimen.

In his second chapter he gives an account

of theories of the evolutionary origin of birds, a fiercely debated subject still. When dealing with controversial issues Feduccia is usually sensible and it is not his fault that at least one of the principal protagonists in this debate has changed sides since the chapter was written. Following chapters deal with the evolution of flight and then with the fossil history (such as it is) and evolution of all the major groups of birds.

Feduccia has been in the forefront of some of the phylogenetic work he describes and much of the book is authoritative. However, particularly in the early chapters, there are some things to take exception to. The layman's naive view of the course of evolution, as a highway from simple organisms to human beings with less important byways branching off, will be reinforced by statements such as "ornithischians [a dinosaur group] were highly specialised herbivores and far too removed from the main line to have given rise to any other major group". It will be similarly reinforced by the use of "Age of Reptiles", "Age of Mammals and Birds" and so on, as in "The cotylosaurs [the earliest reptiles in the author's (outmoded) usage] appeared long before the Age of Reptiles . . .". There are also a few obvious errors — how, for example, can the two views of the original *Archaeopteryx* feather, pictures on p.12, be "main slab" and "counterslab" when they are not mirror images of one another? Also, the patagium (flying membrane) of the living flying lizard genus *Draco* does not stretch "between the forelimb and hindlimb" (p.40): it is supported by elongate ribs alone.

On the whole, though, the book is easy and pleasant to read. The style is clear, concise and interesting, albeit with some infelicities — "Members of this classifica-

tion . . ." (i.e. taxon); "... third alternative . . ." — and the chapters are short. The bibliography is excellent; here Feduccia refuses to "talk down" to the reader and quotes every reference, not just "further reading". However, the illustrations although numerous are often disappointing. They include black and white photographs of specimens and localities, and a number of frankly mediocre original drawings by several hands, the low standard of which is emphasized by some of the splendid and



Reconstruction of *Archaeopteryx* by Rudolf Freund.

famous restorations by Heilmann, Charles Knight and others with which they appear. Many of these cry out for their original colours. Presumably the decision to exclude the use of colour was the publisher's rather than the author's; nevertheless I imagine that it will affect sales adversely. This is a pity as the book is otherwise well produced and certainly deserves an audience. □

Alec Panchen, a vertebrate palaeontologist, is Reader in Vertebrate Zoology at the University of Newcastle upon Tyne.

Gorillas on a tightrope

Alison Jolly

The Natural History of the Gorilla. By A. F. Dixon. Pp.202. ISBN 0-297-77895-1. (Weidenfeld & Nicolson: 1981.) £16.50. To be published in the USA by Columbia University Press.

IF YOU want to know about gorillas, you need this book. Alan Dixon has summarized the present state of knowledge about all aspects of gorillas' biology from anatomy to conservation. Furthermore, he writes in a lucid, factual style which presents the information clearly to anyone from fifth form to professorial status, without either talking down or jazzing up.

The drawings (by Dixon himself) add both clarity and charm.

The book has several strengths. First, it deals equally with the results of laboratory work and results from the field. Dixon's own studies range from laboratory analysis of primate hormones and sexual behaviour, through the testicular atrophy of several adult male silverbacks in zoos (including Guy at London), to making field censuses of gorilla populations. He is thus well placed to appreciate both sides of the story. A second strength is the balance achieved between generalization and data from individual studies. In the chapter "How Close to Man?", he manages to explain quickly many of the bases of primate classification, sum up the evidence which suggests man's close relationship to gorilla and chimpanzee and the more distant relationship of these three African species to orang-utans, and yet include the divergent studies which indicate Asiatic, rather than African, origins for man. Similarly, when he describes wild populations, he includes data showing differences between the three gorilla subspecies, and even between western and eastern Virunga Volcanoes, adding appropriate caution about sample size. In a popular book, it is unusual to avoid generalizations about "the gorilla" meaning ten or twenty individuals on Mount Visoke.

Finally, Dixon covers the whole scope of the literature from Hanno (fifth century BC) up to 1980. This enables him to deal fairly with such experiments as Yerkes' on Congo (1926–1928), which is still one of the most complete accounts of a gorilla's problem-solving, as well as Redshaw's (1978) splendid Piagetian study of the development of infant gorilla intelligence. The conservation account is also up to date, to the death of Digit and others who had their heads and hands chopped off for tourist souvenirs.

The main reservation I have is that Dixon does not pursue his synthesis of detail quite far enough for my taste. I looked for a discussion of the relation of gorilla leaf-eating to metabolism, ranging and social structure, and comparison with the frugivorous orang and chimp. All the pieces are there, but Dixon does not explicitly link them up; similarly, reproductive strategy is split into several different chapters. This is clearly his choice — he gives us the data without fashionable theorizing, but he is better placed than most of us to deal with such theories.

I could also do with slightly more drama in the final chapter on conservation — perhaps some of Dian Fossey's emotion over Digit, or Alan Goodall's encounters with poachers in *The Wandering Gorilla* (Collins, 1979). Instead, Dixon has chosen restraint. Perhaps he is right that estimates of the remaining wild populations and their possible or likely extinction are even more horrifying.

The whole question of how to balance

professional against popular appeal bedevils primate studies. In the circumstances Alan Dixon's summary of the natural history of the gorilla for such a wide audience is laudable. Scholarly balance may seem faint praise until you add that it is balance on a very thin tightrope indeed. □

Alison Jolly is temporarily in the School of Cultural and Community Studies, University of Sussex.

Lure of the shore

J.T. Enright

Intertidal Invertebrates of California. By Robert H. Morris, Donald P. Abbott and Eugene C. Haderlie with 31 contributors. Pp.920. ISBN 0-8047-1045-7. (Stanford University Press: 1980.) \$45 US only.

This is a book for those who find the California shoreline at low tide a place of beauty, wonder and interest, and who want information on the common animals that live there. [Preface].

THE west coast of North America has one of the most diverse temperate-latitude intertidal faunas in the world. In California that fauna provides research material for dozens of marine laboratories and learning material for hundreds of high schools, colleges and universities. *Intertidal Invertebrates of California* is aimed at that limited — some might say, provincial — audience, but it could conceivably also lure naturalists from the entire world to view and study this showplace of diversity. Whether or not it eventually has that sort of incidental success, there is no question about its value and importance for the intended audience: every biologist, professional or amateur, who ventures into the California intertidal zone for the first — or the ten-thousandth — time should have this volume available.

Thanks to a grant from The David and Lucille Packard Foundation, the price is such that most users will be able to purchase their own copies. Even a first glance at the 200 photographic plates — all but three of them in colour, most with four to six separate portraits of the animals — will persuade the prospective buyer that this book is a bargain. The majority of the photographs were taken by Morris, but nearly 50 other naturalists contributed to the handsome illustrations. This portion of the book will be particularly treasured by all who find the "shoreline . . . a place of beauty [and] wonder".

With the photographs as the bait, Abbott and Haderlie, together with 31 other contributors, have provided nearly 700 pages of text, bibliography and index for those "who want information". Some 750 species are treated in detail in the 24 chapters. For each of entry, the text provides the classification of the animal; a morphological description, including

typical size and the features which distinguish that species from those with which it might be confused; and notes on habitat and geographical distribution, often with qualitative comments on abundance. Usually there is additional information on what is known of the animal's natural history. Each entry closes with literature citations, sometimes only one or two, in other cases several dozen. Detailed references are then provided at the end of each chapter.

The introduction suggests that this book was intended as a primary guide for identifying the animals, but this emphasis is, I think, misplaced; it will simply not do as a field manual. Its unusual format and sheer weight (about six pounds) militate against carrying the book along to the shore. Beyond the problems of getting book and animal together (since the reader is, rightly, urged not to bring specimens home with him), there are none of the traditional "keys" for identification; and the coverage is intentionally restricted to the more common and conspicuous elements of the fauna, meaning that using the plates and text for primary identification would be not only time-consuming but often misleading.

For field identification, then, one should instead take along Smith and Carlton's revision of *Light's Manual* (University of California Press, 1980). Thereafter, *Intertidal Invertebrates of California* can be treated as something akin to a biographical dictionary. Knowing, or at least having a strong hunch about the name of the creature, the reader can here confirm that he has, indeed, probably seen this or that member of the community, and then learn some of what the experts know about it, as well as where to find further information.

A conspicuous contrast with a biographical dictionary is that no historian is ever likely, on the basis of some arbitrary legal point, to decide that henceforth Madame Curie will be known only as Marie Skłodowska. Many of my old friends in the intertidal zone, however, regardless of how well established their reputations were, are now assigned names different from those I learned 25 years ago. In some cases (e.g. *Mitella polymerus* and *Aletes squamigerus*) the older names are also included in the index to this volume; but for others (e.g. *Acmaea digitalis* and several congeners, and *Lepidopa myops*) there is no hint about former identity. This is, of

course, not so much a complaint about this book as about the problems inherent in nomenclature, and it can be argued that such changes represent "progress"; but I sometimes wonder. . . .

The neophyte will find the chapter introductions and the detailed text entries in this book most useful; the more advanced student will find the literature citations a good starting point for a research project; everyone will, I think, enormously appreciate the coloured plates. The professional biologist who knows some part of the literature on a particular group or species will no doubt note — as I did —

cases of mistaken conclusions, incomplete or outdated citations, and unevenness in coverage. But it would be petty to emphasize such deficiencies. The central point is that this book is an indispensable addition to the library of all who are interested in the common intertidal invertebrates of the California shoreline — and it contains a very attractive collection of colour portraits for those who will never get there. □

J.T. Enright is Professor of Behavioral Physiology at the Scripps Institution of Oceanography, La Jolla, California.

Old twists to an old tale

Jerry Donohue

The Double Helix: Text, Commentary, Reviews, Original Papers. Edited by Gunther S. Stent. Pp.298. ISBN hbk 0-297-77899-4; ISBN pbk 0-393-95075-1. (Weidenfeld & Nicolson/W.W. Norton: 1981.) Hbk £10; pbk \$5.95.

AN OLD proverb, as given by Gibbon, tells us that "happy is the nation that has no history". If this applies to substances, then DNA, surely, is the unhappiest of molecules, for it is the subject of innumerable biographies.

This new book contains: the complete *The Double Helix*; reprints of 12 reviews of it from 1968 and discussions of these by Gunther Stent; reprints of three papers from *Nature* by James Watson, Francis Crick and Aaron Klug; three letters to *Science* from 1968 by Max Perutz, M.H.F. Wilkins and Watson (see below); reprints of six papers from 1953–54 describing the original DNA work done in Cambridge and King's; and a name index the usefulness of which is somewhat marred by careless inaccuracies.

Should the fact that only about 5% of the contents of this new edition is previously unpublished material be cause for discontent? Probably not, as it brings together *The Double Helix* and some of the important writings it spawned. There are, however, other causes for discontent which are hardly contentious. The original text of *The Double Helix*, including the photographs and diagrams, ran to 233 pages; in this reprint these occupy 135 pages. This drastic compression was accomplished by abandoning the elegant format of starting each chapter on a new page; by taking the plates of the first (Signet, 1969) paperback edition, which was itself a compression of the original, and compressing it still further by putting more lines per page; and by reducing the sizes of the photographs, some by more than 50%, so that the individuals in the group photographs are virtually unrecognizable. Also reduced in size is the photographic reproduction of an historic

letter of Watson to Delbrück which is legible only with great difficulty and, in some places, totally undecipherable, a circumstance not caused by Watson's crabbed hand. Surely a book as important as this deserves a better production.

Something very mysterious should also be pointed out. In the diagram on p.111 the enol structure for thymine incorrectly has an oxygen atom where there should be a hydrogen atom. The same diagram on p.191 of the hardback first edition of *The Double Helix* is correct, and identical in every other respect, right down to the pattern of dots inside the rings of the bases. Now, admittedly, some of the early difficulties in cracking the DNA structure concerned the keto-enol business, but what sinister force, in Alexander Haig's terminology, took it upon itself to change a correct structure to an erroneous one?

There are several misconceptions and misreportings that ought to be dispelled before remarking on a few of the reviews included in this new volume.

● *The Myth of the Invention of Base Pairing.* Throughout the reviews runs the thread of the notion that Watson and Crick would never have discovered their structure had it not been for "rules" of Erwin Chargaff that in DNA (molar amounts of) cytosine = guanine and thymine = adenine. However, Chargaff himself had not yet referred to this relationship as rules, and, indeed, his own published analytical data gave merely the merest hint of these regularities, and he himself averred the possibility that the regularities might be accidental. Stent (pp. 168–169) has more to say on this.

The fault for these misconceptions must be squarely on the shoulders of Watson. We find him writing (p.75) "The moment [Spring 1952] was thus appropriate to think seriously about some curious regularities in DNA chemistry first observed by Chargaff", and then subsequently cheerfully disregarding them, and building a like-with-like, non-complementary, model. In addition, he had been assured by his friend, the biochemist Markham, that if

Cold Spring Harbor Laboratory have recently published the proceedings of the symposium on viral oncogenes which was held last year. The two-volume set of proceedings, comprising Vol. XLIV in the series *Symposia on Quantitative Biology*, is available from CSHL, Fulfillment Department, PO Box 100, Cold Spring Harbor, NY 11724, or through booksellers, price \$130 (\$156 outside the USA).

Review of James Watson's *The Double Helix* by J. Field

Hear the song of how the spiral
Complex, twisted, double, chiral
Was discovered. Its construction
Following some shrewd deduction
Was accomplished. From all sides "Oh
Tell it as it was" they cried so
Then he took his pen and paper
And composed the helix caper
Writing of the Cambridge popsy
Turned the world of science topsy
Turvy like had not before been
Done. The story is no more than
How the structure was discovered
Secret of the gene uncovered
Biological prediction
Better than a work of fiction
Stylistically breezy
Everything appeared made easy
And it was. The bases' pairing
Found at last — a small red herring
Notwithstanding. How he found them

And the hydrogens that bound them
Was a stroke more accidental
Than a work experimental
Chemistry was not his calling
He had read *one* book by Pauling
Models of the bases he'd made
With them on his desk top he'd played
Shuffling them and putting like with
Like he made no lucky strike with
Guanine, thymine, adenine. A
Letter came from Pasadena
Horrors! then the day was won for
Pauling's model was quite done for
Extra atoms he had set in
Where no atoms should be let in
With some phrases less than modest
Pauling's model's called the oddest
Back now to the basic pairing
How's the structure building faring?
Faster, faster goes the race, then
Everything falls into place when

Tautomers which nature chooses
Are the ones our author uses
Two chains (but here we can't be sure)
Plus this extra structure feature:
Pauling's outside bases inside
Pauling's inside phosphates outside
Now they all were quite ecstatic
Soon became they quite dogmatic
Having nature's secret later
Checked against the x-ray data
This was difficult, for Rosie
Found our scientist rather noseey
Bragg compares the book with Pepys. Is
This the verdict of Maurice? His
Thoughts alas we cannot gainsay
Trumpet blowing's not his forte
Tales like this do have a moral
Whether printed whether oral
Find the right man to advise you
Then you'll get a Nobel prize too.

Chargaff said that guanine equalled cytosine, he, Markham, was equally certain that it did not, because Chargaff's experimental methods inevitably underestimated the true amount of cytosine. On numerous other occasions Watson refers to Chargaff's "rules".

It is, on the other hand, abundantly clear from the narrative that these "rules" had nothing whatsoever to do with the construction of the DNA model. The model came first, and made the rules, not the other way around.

● *The Myth of the Undercover Agent.* In some of the reviews selected by Stent the role which Peter Pauling played is displayed as not quite what it should have been. Thus Watson, according to Merton (p.215), reports "with temerity and self-mocking wit the occasions on which Linus Pauling's son, Peter, then a student at Cambridge, became a prime source of information about what his father was up to". Morrison (p.176) says that the idea for the double helix came "... out of hearing from the young American expert upstairs [actually, I was not upstairs at all, but in the same office as Bob Parrish, Watson, Crick and Peter] that the textbook was probably wrong, or out of rumors about what Pauling was thinking — relayed by his son Peter at Cambridge". Lear (p.195) goes considerably further in stating that Watson "discloses how he used his young friend Peter Pauling to spy on Pauling's brilliant father, Linus". And Lwoff (p.227) observes that "Peter Pauling works in the Cavendish, receives detailed letters from Pasadena and informs his colleagues of the evolution of his father's work".

On careful reading of *The Double Helix* we find that Watson says nothing of the sort. Consult, if you will, the name index for the citations to verify this, and come to the conclusion that these statements and innuendoes in the reviews are pure fiction. So much for spies!

● *The Myth of the Race.* Another

thread, nay, hawser, which twines through the reviews, is the "breathless race" between Watson and Crick on the one hand, and Linus Pauling on the other, towards that legendary ceremony, established by the dynamite tycoon, in the gloomy, sunless Stockholm of December. The same strand snakes its way through *The Double Helix*. How, though, can there be a race with only one team and no antagonist? At this stage (1952–1953), Pauling, whose reputation as a scientist was unequalled, had well over 200 publications under his belt, including a dozen or more books universally appreciated and widely read. His ticket to Sweden was long overdue. (He did, in fact, make this journey the very next year in 1954 to collect his first Nobel prize.) Was he racing Watson? Certainly not. To Pauling, in Peter's view (*New Scientist*, May 31 1973, p.558), DNA was just another chemical, intrinsically equal in interest to, say, sodium chloride; both presented interesting structural problems, so why race with DNA?

● *The Myth of the Filched Photographs and the Dipped Data.* The fourth and final myth concerns the X-ray data, and whether Watson and Crick came by them by unorthodox methods. Among the reviewers we find Sinsheimer referring to "cadged data" and a "privileged report", Lear to a "clandestine communication channel" and supposedly "confidential data", and Merton to "an odd expedient for gaining access to badly needed information". Lwoff and Chargaff (in his review, not included in this volume) quote a crucial paragraph from *The Double Helix* which led to both of them to comment on ethics. The same passage is reproduced by Perutz, the supposedly clandestine channel, in his letter, pp.207–210, in which, as well in the letters by Wilkins and Watson which follow, it is made abundantly clear that the report was not confidential, and the data had already been

disclosed in open seminars.

As a matter of fact, an X-ray diffraction photograph which, although not of the high quality of those produced by the two groups at King's, clearly shows the characteristic helical pattern of the B structure, had been published by Astbury years earlier, so all the fuss over stolen data was really unnecessary.

Five of the reviews of *The Double Helix* included here were previously reviewed by Gunther Stent, the editor. The article by Chargaff, that *vieux terrible* of the DNA saga, although not reprinted because he withheld permission, is reviewed. (Anyone wishing to read the original will find it in *Science* 29 March 1968; p.1448). Of the seven new reviews, one of them warrants comment, that by the bipseudonymous "F.X.S.", originally "F.R.S.". This is not a review at all, but verbal virtuosity by a polylingual sesquipedalianist, and why Stent chose to include it is an enigma, as it adds little insight into DNA or those connected with it. F.X.S. is coy about his identity, but really, who cares? His piece is worth reading, but hardly reprinting.

There are numerous other important reviews not included by Stent — by Lord Todd in *Chemistry in Britain*, Sokolov in *Newsweek*, Bernstein in *The New Yorker*, Maddox in *Nature* and so forth. Another not selected by Stent, by J. Field, appeared in the *Journal of Irreproducible Results*, (Vol.17, December 1968, p.53). Because it expresses as well as anything the flavour of Watson's book, it is reproduced above (with the permission of the editor).

Anyone who hasn't yet read *The Double Helix* should certainly do so, but, even if you own a copy, the additional material in the present volume amply justifies its purchase. □

Jerry Donohue is Rhodes-Thompson Professor of Chemistry at the University of Pennsylvania. He spent the academic year 1952–1953 as a Guggenheim Fellow at the Cavendish Laboratory, University of Cambridge.

Schumacher reconsidered

Eric Ashby

Small is Possible. By George McRobie. Pp.331. ISBN hbk 0-224-01858-2; ISBN pbk 0-06-090694-4. (Jonathan Cape/Harper & Row: 1981.) Hbk £7.95; pbk \$5.95.

THERE is a mysterious jackpot logic about best-selling books on the environment. Why has Vogt's *Road to Survival* sunk without trace while Leopold's *Sand County Almanac* is quoted and requoted? Both were published a generation ago; both are elegantly written. Why is Thoreau, who liked his wilderness to be as tame as a tabby cat, still regarded as a conservationist's guru? And why is Schumacher's *Small is Beautiful* treated as a seminal tract in self-sufficiency? Surely Tolstoy had similar ideas and there were Tolstoyan communities (well described by W.H.G. Armytage in *Heavens Below*, published 20 years ago) in Britain in the 1890s.

Small is Beautiful, I'm told, still sells well, thanks (ironically) to the large-scale techniques of book production and marketing. It is seductively written and full of wisdom. But wisdom, alas! is only a minor ingredient in practical politics. As devotional reading, Schumacher's book is still popular; but as a blueprint for a new economic policy in the manifesto of a political party (except, perhaps, the misnamed ecology party) it's not a starter. Not only in the capitalist world but in the communist world too, small may well be beautiful, but it is the giant enterprise that provides energy, packaged foods in the supermarket, communications, transport, the dish-washer in the kitchen and the car in the garage. The communes in China are an exception to this generalization, but will they survive as China becomes a great industrial power? For Britain, at any rate, it does not seem unduly cynical to assume that if a choice had to be made at the polls between an economic life style which was, in Schumacher's sense, beautiful, and one which gave the greatest material comfort to the greatest number of people, then beauty would lose out.

So the stock comment on *Small is Beautiful* is "beautiful, no doubt; practicable, no". Therefore when a sequel appears with the title *Small is Possible*, it should be accepted as a challenge to re-examine Schumacher's thesis. It is in that spirit that I read the book.

George McRobie, the author, was Schumacher's assistant and co-founder of the Intermediate Technology Development Group, ITDG (one of the many — far too many — acronyms in the book). He does not pretend to write with the felicity of Schumacher but he doggedly assembles evidence to validate the book's title. The book is, for the most part, an annotated catalogue of enterprises in intermediate or

"appropriate" technology which are consistent with Schumacher's thesis, namely that conventional capital-intensive projects are a disaster in Third World countries and may well be the signpost toward disaster in industrialized countries; and — this is the message of the book — that there are alternative small people-intensive projects which are viable. McRobie gives a list, with addresses, of nearly 100 organizations concerned with appropriate technology engaged in hundreds of little experiments. The evidence persuades me (at any rate) that there is a minority movement to reject the life-style of what used to be called the military-industrial-complex; not on political grounds but on what may become a much firmer foundation: moral grounds. It is no more than a sprinkling of enthusiasts (some of them cranks) in the interstices between super-highways, power stations, acres of petrochemical works and refineries; uncoordinated, fragmented. But something interesting is going on, and if the fragments coalesce a new life style might emerge.



The late E.F. Schumacher, prophet of self-sufficiency.

What is going on? Here are a few samples taken from McRobie's book.

He deals first with the most urgent and relevant application of Schumacher's thesis: the design of equipment for agriculture and industry in Third World countries. The principle is simple. In countries with surplus manpower, and without technical sophistication, capital investment per worker has to be as low as possible. When a British trade mission went to Nigeria, Schumacher's group (ITDG) put into their hands a stencilled *Guide to Small-scale Equipment for Rural Development*, "of animal and man-operated agricultural equipment obtainable from little hole-and-corner firms in Britain". Not at all what the mission intended to propagate in Nigeria; but (said Schumacher) "This thing was torn out of

their hands by people who said: For the first time you bring us something of interest, instead of the usual glossy catalogues". That was 15 years ago. Since then the ITDG has had an impressive record of service to Third World countries. A building panel set up, in five African countries, units for making bricks at a capital cost of £400 per workplace, compared with £40,000 per workplace in a large modern brickworks; and it invented and put into production a fibre-reinforced cement for roofing at a fraction of the cost of the normal commercial product. A water group published handbooks on the making of pumps, wells and low-cost irrigation systems. A farming group has designed and put into production (in Zambia and Nigeria) workshops for making machines for weed-control, groundnut harvesting and fibre stripping. A transport unit based on the Engineering Department at Oxford has designed the Oxtrike: a cycle rickshaw which can be made in Third World countries. Windmills in Ethiopia, water pumps in Nepal, mini-hydro generators — all locally made — replace the imported equipment which cannot be kept in good repair without skilled technicians. And — most imaginative of all — the Rural Health Panel has, under the skilled guidance of Dr Katherine Elliott, promoted the training of auxiliaries in health care, in collaboration with the World Health Organization.

These efforts may seem to be trivial crumbs to offer the Third World when set beside the massive expenditures on steel mills for India or on the Volta Dam for Ghana; too trivial for all the effort involved. But they have an importance which cannot be quantified: they may, with luck, convince the Third World that it does not want to ape Europe and America in its patterns of society. And it may help to repair some of the damage we have done in exporting the wrong kinds of goods and education to the Third World. Thus, when we exported medical education to the university colleges of tropical Africa we arranged (with the noblest of motives) that its quality and content would be indistinguishable from that in a London medical school. The surgical staff, training doctors skilled to remove the gallstones of cabinet ministers in the capital city, regarded with impatience the activities of an African professor of social medicine, who put his efforts into the holding of competitions in the villages for the construction of earth closets. Anyone who dared to suggest that the prime need was for medical auxiliaries (as the Soviet Union and China had) was scarified with criticism. The universities were suspicious of any subject in the curriculum which would not have been appropriate in London. One encouraging message in McRobie's book is that this technological snobbery is diminishing. African universities are becoming relevant to African problems. Countries where Small has been

Practicable (if not Beautiful) for millennia may still have a chance to escape some of the social perils of modern technology.

What about the industrial West, where the proudest boast of technologists is that they save manpower at the cost of capital investment per workplace? It would be stupid to dismiss the enormous social benefits of mass production. For every one man whose life is dissipated on a production line in British Leyland there must be hundreds of families whose lives have been liberated by the possession of a car. Holidays, picnics, visits to friends and relations, life in a suburb and a short drive to the office: all these are incomparably easier to secure than they were for the so-called lower classes when I was a boy. While there was full employment, and material progress brought with it some improvement in the quality of life, the achievements of the industrial West were welcomed. For the first generation owner of a car and a refrigerator and an air ticket to Majorca, the warnings of Schumacher were dismissed as neurotic.

It is a different story today. Thanks to education and the media, the unemployed are more articulate than they were in the 1930s. They do not equate unemployment with leisure, even though they are paid to be unemployed. To be without work is as dangerous as to be without love. So, suddenly, the application of Schumacher's ideas (and not just Schumacher's: they are the ideas of William Morris and Tolstoy in modern dress) becomes relevant and politically important. McRobie describes how hundreds of small groups of people in America, Canada and (on a smaller scale) in Britain have detached themselves from the conventional pattern of society, to seek a pattern which does not have the formula "economies of scale" woven into the design. In America some of the experiments have official support. There is a National Center for Appropriate Technology, inspired by Senator Mike Mansfield. There is the California Office of Appropriate Technology, started by order from Governor Jerry Brown, with a budget of \$900,000, encouraging the use of solar energy, and designing small-scale sewage systems and techniques for aquaculture. In the Bronx, some of which is a waste of derelict city lots, an Institute for Local Self Reliance has mobilized local labour to cultivate gardens and rehabilitate decayed property. And in Canada — in many ways the most enlightened of Western nations — there is in the ugly mining town of Sudbury a project called "Sudbury 2001" which, with a provincial government grant of \$600,000, is preparing some of the workforce for the inevitable redundancies they see ahead.

McRobie's chapter on the "alternatives" movement in Britain is less encouraging. It seems (perhaps there is another side to the story) that neither governments nor trade unions have taken much interest in the possibilities of small

collectives. The experiences of the Kirkby Manufacturing and Engineering Ltd (KME) and the *Scottish Daily News* do not generate much confidence. Of course it is sheer fantasy to imagine that ICI, BP and the Central Electricity Generating Board will dissolve in a flood of little collectives making their own paints, heating their farms with home-made methane, using windmills knocked together by the local blacksmith. But the unemployed displaced

by a combination of world diseconomy and new technology are going to demand work. If they do not get work, society as we know it may be shattered. If we have to invent work not to produce things for profit but to satisfy a deep genetically coded human need, then perhaps Schumacher's ideas are worth another thought. □

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge.

Science fiction as Marxism

James Gorman

Robert A. Heinlein: America as Science Fiction. By H. Bruce Franklin. Pp.232. ISBN hbk 0-1950-2746-9; ISBN pbk 0-1950-2747-7. (Oxford University Press: 1981.) Hbk \$18.95, £10; pbk \$4.95, £3.95.

READING H. Bruce Franklin on Robert Heinlein is a bit like reading an analysis of a sexual farce by a stiff-necked, bible-bound fundamentalist who can never get his mind off the fact that sex is, after all, sinful. Franklin has produced a painstakingly detailed Marxist analysis of what, at its best, is flashy escapist fiction. Heinlein's heroes, his cynicism and sentimentalism, his taste for adventure, all his tales of space exploration, war, sexual fantasy, time travel and immortality become a rather dreary reflection of the contradictions of capitalism in the United States.

Franklin leaves the science in the fiction more or less alone; this omission does no disservice to Heinlein. Science only provides him with the authority by which reality is breached — it is the excuse for the reader to believe in the fantasy. When it suits his purpose, Heinlein is not averse to using magic in its place.

Heinlein is at least as fond of social theory as he is of science, and his books are full of expostulation on one half-baked political philosophy or another. And, as Franklin points out, he does reflect his times. In the 1950s he produced anti-communist novels; in the 1970s he explored narcissistic sexual excess in some of the worst writing of his career. And, in 1961, at the beginning of one of America's most notorious decades, he produced what is probably his most popular book, *Stranger in a Strange Land*. In it a young human being, raised on Mars, by Martians, comes to Earth, views us all through alien, though human eyes, and finally finds a religion that is part universal sexual love, part hucksterism, part anarchism and part violent elitism. The gentle, loving hero is in the habit of making his enemies, and other undesirables, disappear by a wink of the mind, a trick he learned on Mars.

This is, indeed, very close to "America as Science Fiction". But when Franklin presses his point to argue that the hero "embodies our most infantile fantasies and

the central goal of bourgeois ideology — the unfettered freedom of the individual will", he is less convincing. Indeed, when, throughout the book, he holds up Heinlein as a mirror that reflects the nature of capitalist society, he undermines his case by two unfortunate omissions — he leaves out Heinlein, and he makes his analysis on the basis of unargued and undemonstrated assumptions.

Where indeed is Heinlein in all of this? Franklin gives us only a thumbnail biography, which concentrates more on the economic conditions in the country when the author grew up than it does on Heinlein himself. And where is Heinlein's imagination, his writing skill, his ability to create compelling stories? A writer of fiction is less a mirror than a lens, that refracts rather than reflects the light of his culture, bending it with his own idiosyncrasies. In Franklin's formula the curve of Heinlein's lens is missing.

Also missing is some back-up for his ready-made Marxism. It may well be true, as he notes in passing, that the conditions of the two main groups of workers in America before World War II were "the debt peonage characteristic of the majority of Black rural workers and the wage slavery of mine and factory labor", and that "class struggle . . . has constituted actual human history", but his case would be stronger if he demonstrated his points rather than asserting them. Nor are these isolated examples. He tells us, for instance, that "If the will is free to do anything it wishes, the will is free from the apparent laws of the physical universe and also free from the apparent laws of human social development". He doesn't tell us what either set of laws are, perhaps assuming that the precepts of dialectical materialism are as well known and widely accepted as gravity.

What Franklin's book could use, finally, is a bit more of an internal dialectic. If we were to see the clash of culture and individual in Heinlein's work, we might come away with a richer understanding of the author, and of America. □

James Gorman is a staff writer on *Discover* magazine.

Spring Books

Gel Electrophoresis of Proteins: a practical approach

Edited by D. Rickwood and B. Hames

A laboratory aid for all those wishing to separate proteins by gel electrophoresis. Contains detailed recipes and step-by-step instructions on the procedures involved in all the different types of electrophoretic separation.

0 904 147 22 3 (soft) £7.50/US \$18.00
300 pp (approx.)

Rapid and Automated Methods in Microbiology and Immunology: a bibliography 1976-1980

A unique reference tool to help you find the simplest method or technique best-suited for your laboratory. Contains over 3000 citations on practically all the rapid and automated techniques introduced in the field during the last five years.

0 904 147 19 3 (soft) £15.00/US \$36.00
200 pp (approx.)

Tooth Surface Interactions and Preventive Dentistry

Edited by G. Rolla, T. Sonju and G. Embery

Proceedings of a workshop held December 1980 in Oslo and part of a series devoted to current aspects of dental research concerned with the interactions occurring at the tooth surface. Particular emphasis is directed at the role of inorganic ions, especially stannous fluoride, in influencing the metabolism of oral bacteria and their adherence properties.

0 904 147 29 0 (soft)
£15.00/US \$36.00 217 pp, 40 illus.

Immunology of the Eye Proceedings of three workshops sponsored by the National Eye Institute, USA

Workshop I: Immunogenetics and Transplantation Immunity

0 904 147 25 8 (soft)
£10.00/US \$25.00 300 pp

Workshop II: Autoimmune Phenomena and Ocular Disorders

0 917 000 07 2 (soft)
£10.00/US \$25.00 200 pp (approx.)

Workshop III: Infection, Inflammation and Allergy 0 917 000 08 0 (soft)
£10.00/US \$25.00 300 pp (approx.)

Prepaid orders and requests for further details to:

Information Retrieval

1 Abbey Street, Eynsham,
Oxford OX8 1JJ, England
1911 Jefferson Davis Highway,
Arlington, VA 22202, USA

Circle No.31 on Reader Enquiry Card.

Colourful tale of two Voyagers

J. Veverka

Voyage to Jupiter. By David Morrison and Jane Samz. Pp.199. (National Aeronautics and Space Administration: 1980.) Pbk \$7.50 US only.

NASA-sponsored books describing flight missions to various planets are not hard to find. Unfortunately, good ones are as scarce as impact craters on the surface of Io. Typically, such books amount to little more than collections of pictures intermingled with spacecraft details and other mission minutiae. Science gets mentioned only in passing and at so superficial a level that even the casual reader feels short-changed.

Voyage to Jupiter is not such a book. It is a beautifully illustrated, vivid account of the two historic Voyager encounters with Jupiter in 1979, as well as a factual summary of the initial scientific results. Everyone should find it interesting; even the sophisticated reader will find it informative.

In part, the successful outcome may result from the combination of talents of the two authors: Samz, a science writer, and Morrison, a planetary scientist and a member of the Voyager Imaging Team who was personally involved in many of the discoveries described in this book. Certainly, this volume does contain details of the spacecraft construction, the mission profile and the requisite pictures of the major participants, but interwoven with all this is real, substantial science. Voyager's many discoveries — rings, volcanoes, lightning, new satellites and so on — are not merely catalogued chronologically, but are tied into a comprehensive new picture of Jupiter and its satellites. Without question, *Voyage to Jupiter* is more than a NASA picture book, but it is certainly that too. The Voyagers are the first planetary spacecraft to have sent back large quantities of true colour data, and it would be difficult to imagine a better subject for colour imaging than the jovian system. Many of the best of these pictures are reproduced admirably in this volume.

The book opens with a pair of chapters summarizing our pre-Voyager knowledge of the Jupiter system based on three centuries of telescopic observations and the initial reconnaissance by Pioneers 10 and 11 in 1973-1974. Next, the authors discuss the spacecraft, the scientific work that its eleven instruments were designed to carry out and some of the individuals responsible for these experiments. Then come three chapters giving a day-by-day account of the two encounters. Here the reader can share some of the excitement experienced by those who were fortunate enough to be in Pasadena during these two historic events when major discoveries were being made almost hourly. Two of the remaining three chapters aptly summarize our

current, post-Voyager, view of the planet and its satellites, and the volume ends with a discussion of future missions to Jupiter. Sadly, this chapter must be very short. At present, the only mission planned is NASA's Project Galileo, which will investigate the jovian system in the latter half of this decade. For a while at least, our view of Jupiter and its satellites will remain that relayed to us by Voyager in 1979.

Voyage to Jupiter — an excellent little book — sets a high standard for future NASA publications of this sort. In these days of \$30-\$40 science books, it is an absolute bargain which I heartily recommend. □

J. Veverka is at the Laboratory for Planetary Studies, Cornell University, New York.

Here comes Halley

David W. Hughes

The Comet is Coming! The Feverish Legacy of Mr Halley. By Nigel Calder. Pp.160. ISBN 0-563-17859-0. (BBC: 1980.) £8.75. To be published in the USA by Viking, \$12.95.

"EVERY seventy-six years or so Halley's comet punctuates history like an exclamation mark in the sky". It was last seen in 1910 and it is due back soon. To be precise it will be at perihelion, the point in its orbit closest to the Sun, on 9 February 1986. The preparation for this momentous event has already started. The European Space Agency is busily organizing the Giotto spacecraft which will pass as close as possible to the cometary nucleus. The International Halley Watch is being supervised by the Jet Propulsion Laboratory in Pasadena and this will co-ordinate all the ground-based and near-space observations of the comet. The public are getting ready



Edmund Halley, Astronomer Royal from 1720 to 1742.

Hulton Picture Library

too. Brian Marsden once said "To the man in the street the Solar System consists of Mars, the rings of Saturn and Halley's comet". Comets fascinate people and the next few years will see a flood of books, television programmes and newspaper articles on the subject.

Nigel Calder and the BBC are in the vanguard and *The Comet is Coming!* is the book of the television programme that is to be shown first on BBC2 in mid-1981. The book is excellent. Calder has a beautiful turn of phrase and often had me chuckling. The text is non-mathematical and infectious readability. The illustrations have been carefully selected and many are new to me — not for Calder is the regurgitation of the old faithful. The style is journalistic but you are never left in doubt that the author is a sensible scientist.



Portent of a new King? Detail from the Bayeux tapestry showing the appearance of Halley's comet.

Halley's comet and its like are ostentatious and elusive and because of this they create a no-man's-land of speculation where crazy ideas may be fantasies or brilliant insights. This book is a sure guide to this no-man's-land and underlines the joys, mysteries and frustrations of cometary science. We read of the attitude of the superstitious who regard comets as telegrams from the Gods — disorders in heaven importing changes of times and states. The role of Edmund Halley is stressed, especially his work on orbits. The physical and chemical nature of comets and the suggestions as to their origin are reviewed, as are the problems of cometary discovery and observation. Fred Hoyle and Chandra Wickramasinghe's theory that comets could have brought life to Earth and are even today scattering disease over the globe through the influx of dust is discussed. Calder examines the consequences of a collision between the Earth and a comet; and special reference is obviously given to the demise of the dinosaurs about 65 million years ago!

The book is well referenced, well indexed and I recommend it strongly. □

David W. Hughes is a Lecturer in Physics and Astronomy at the University of Sheffield.

Star-spangled dream

Brian O'Leary

Bound For the Stars. By Saul J. Adelman and Benjamin Adelman. Pp.335. ISBN hbk 0-13-080390-1; ISBN pbk 0-13-080382-0. (Prentice-Hall: 1981.) Hbk \$17.95, £11.65; pbk \$8.95, £5.80.

"WE ARE too great a nation to limit ourselves to small dreams" said President Reagan in his inaugural address. Amongst the budget slashing that has followed, it is difficult to imagine just what dreams Mr Reagan has in mind.

One thing is clear: there is no shortage of big dreams in space. Sound engineering concepts wait for their political and commercial opportunity. Once they are triggered, we can expect to witness an explosive renaissance in the high frontier. The question is not whether it will or should happen; it is when, how and by whom. Soviet-American political and military pressures mean we are in space to stay.

Different writers, however, have different goals. Gerard K. O'Neill, in *The High Frontier*, argues the case for satellite solar power stations and space habitats manufactured in space from lunar materials. Carl Sagan, in *Cosmos*, describes the excitement in expanding scientific programmes in modern astronomy, the search for extraterrestrial intelligence and unmanned planetary exploration. In *The Fertile Stars*, I show how asteroid mining and retrieval could provide the key to expanding the limits to growth of the world supply of energy, food and raw materials, to pave the way to space colonies and manned planetary exploration. Saul and Benjamin Adelman have a rather less accessible dream to offer; *Bound for the Stars* champions interstellar travel as the primary goal for future space activity.

There is little question that NASA's lack of a well-defined objective has hurt the American space program immensely. The lunar-landing commitment made by President Kennedy in 1961 (and completed eight short years later) was the last grand goal of space. We are now seeing the final vestiges of Apollo momentum petering out — the Voyager Saturn fly-bys will be the last American planetary probes to return new data for at least another five years.

Apart from the space shuttle, the US civilian space programme is virtually dead, while the Soviets, Europeans and Japanese move steadily forward. For example, all three plan a Halley's comet mission during its once-in-a-lifetime appearance in 1985 and 1986; budget cuts have grounded an American Halley probe. The Adelmans say these retrenchments can be ascribed to shortsighted year-to-year planning in the US federal government, and contrast this with the project-approval system of the European Space Agency. They stress the importance of changing NASA's charter to protect it from the lack of continuity it

suffers from the annual budget cycle.

The best feature of the book is its wealth of detail on concepts of interplanetary and interstellar travel. The authors trace the history of the development of propulsion systems and navigational techniques with skill and perceptiveness. Most of the concepts have not flown. They describe, for example, the ill-fated Orion Project (in which exploding atomic bombs propel a large spaceship) as a "science orphan no one wanted to adopt".

The Adelmans' stand that interstellar travel should be the basic unifying aim for future space activity is controversial. Because starships are well beyond our current technological capability, many decision-makers will not accept such a heady goal. By contrast, Apollo was — and manned Mars landings, asteroid mining, space industrialization, space colonies and satellite solar power are — feasible projects which could use today's technology.

The first few chapters of the book are a smorgasbord of what appear to be afterthoughts. The authors skip rapidly from the origin of life to the search for extraterrestrial intelligence to a survey of the Solar System, back to life in space, plans for biological experiments aboard Spacelab, controlled ecologies, Solar System travel, space colonies and so forth.

These sections are weak, covering the subjects incoherently with misleading and biased remarks: for example, the belief that all life must be carbon-based and depend on water (see *Life Beyond Earth* by Gerald Feinberg and Robert Shapiro for a different view); "Dr Carl Sagan of Cornell University believes that about a million civilizations exist in the Galaxy on the same or more advanced level than Earth civilization" (p.17, too strong); "The galactic cosmic rays from outside the solar system are not a serious hazard, even for voyages lasting more than a year" (p.41, not true). In addition, they show a strong bias against satellite solar power and for nuclear power (pp.80–85) which, in the United States, might prove politically unacceptable.

The real strength of the book, then, is in its thorough discussion of the prospects of interstellar travel. I have been looking for a handy reference on the subject, and this is most certainly the best available. I applaud the Adelmans' remarks about the importance of keeping the space programme alive and that we need a goal to work toward.

I believe political and military pressures are the foot in the door, economics will become the driver (space industrialization) and the thorough exploration of the Universe will be the ultimate and most fulfilling by-product.

In spite of all the budget-cutting and pessimism, we have some ideas. The Adelmans' treatment of interstellar travel in this new book is one of them. □

Brian O'Leary is a former scientist-astronaut, and author of *The Fertile Stars*, to be published in June.

The pluviiculturists

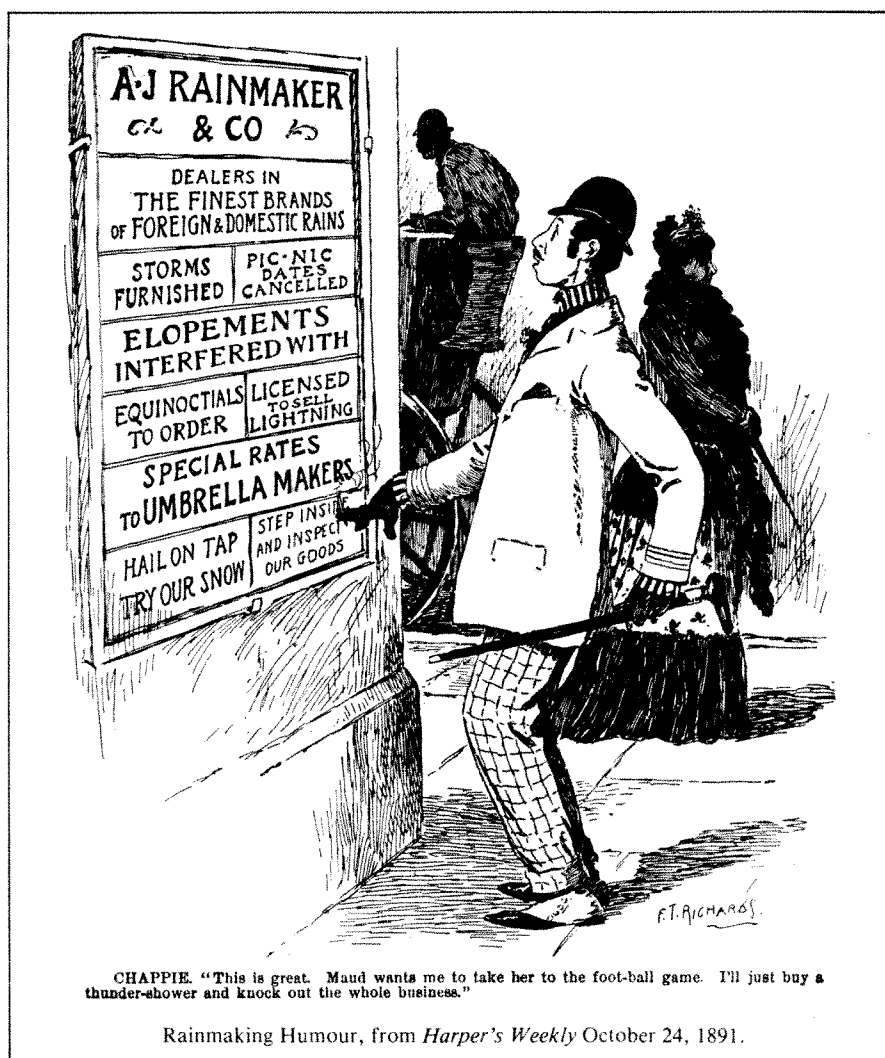
Richard W. Katz

The Rainmakers: American "Pluvi-culture" to World War II. By Clark C. Spence. Pp.185 ISBN 0-8032-4117-8. (University of Nebraska Press: 1980.) \$15.95, £9.60.

In 1930, "Dr" George Ambrosius Immanuel Morrison Sykes was employed by the Westchester Racing Association to ensure sunny weather at Belmont Park racetrack in New York City. Although not a meteorologist, Sykes called himself a "meteorologist", that is, someone who can actually control the weather. Besides suppressing rainfall, Sykes' Weather Control Bureau also claimed to be able to prevent snow, hail, frost and hot weather. The Bureau could, on the other hand, generate rainfall if so desired. To perform these seemingly impossible feats, Sykes relied upon a bizarre assortment of equipment, ranging from a toy propeller to a tub of smelly chemicals. His efforts to guarantee sunshine at Belmont Park, needless to say, met with only mixed success.

Although the putative beginning of scientific weather modification dates from 1946, with Vincent Schaefer's successful seeding of clouds using dry ice, Clark Spence's entertaining book describes the largely forgotten history of earlier, less-scientific rainmaking attempts in America. David Jordan coined the term "pluvi-culture" to refer to the endless rainmaking schemes that were inevitably proposed upon the occurrence of drought. Most of these "pluviiculturists" were not nearly as outlandish as "Dr" Sykes. They included James Espy, a true scientist, whose theories concerning a relationship between fire and rainfall were respected by the scientific community. The most popular rainmaker was probably Charles Hatfield (the "miracle man"), a former sewing machine salesman, who used the "smell maker" technique involving the release of chemicals into the atmosphere. Reputedly responsible for the great flood of 1916 in San Diego ("Hatfield's flood"), he has become a part of the folklore of southern California.

The recipe for success as a rainmaker was simple. Above all, it required being a convincing salesman to take advantage of the public's gullibility. When faced with drought, farmers were desperate and



willing to try almost anything. This "drought psychology" made the "no rain, no pay" schemes offered by the rainmakers highly attractive. Aided by weather forecasts and information on local meteorological patterns, the pluviiculturists depended on occasional coincidences with the occurrence of natural precipitation to demonstrate the viability of their rainmaking operations. Plausible excuses could always be dreamed up in case of failure. Sensationalized journalism helped too, publicizing apparent successes, while not bothering to mention failures. The scientific community, including the US Weather Bureau, remained steadfast in opposition to nearly every rainmaking scheme, but to no avail; scientific expertise was no match for expert salesmanship.

The rainmakers took advantage of several common fallacies, all somewhat statistical in nature. The *post hoc, ergo propter hoc* fallacy arose in connection with the coincidence between rainmaking operations and subsequent rainfall. Pluviiculturists always took credit for any rainfall, even if it occurred several days after their operations or only at distant locations. Another fallacy concerned the often misused "law of averages". Rainmakers were well aware of the odds for rain, correctly realizing that it was sure

to occur eventually. That the public was confused about the law of the averages is no wonder; even Spence is apparently confused about it. He suggests that the task of a rainmaker was made easier, because an extended drought increased the likelihood for rain.

It is impossible to read Spence's book without being compelled to make comparisons with more recent weather modification activities. Some striking similarities are present. Overstatement continues to be employed to sell rainmaking, with one current euphemism for such activities being "weather resources management". Journalists' reporting of the subject remains somewhat misleading or inaccurate, and the public tends to accept the same fallacious reasoning. As always, interest in pluviiculture is revived whenever a drought occurs. To justify government funding of weather modification programmes, politicians rely to some extent on the same "it's worth a try" or "it can't hurt" arguments. If Spence's history of American rainmaking were extended to the present, how different would it be? □

Richard W. Katz is in the Department of Atmospheric Sciences at Oregon State University, Corvallis.

Two psychology textbooks recently reviewed in *Nature* (289, 710; 1981), *Behavioral Neuroscience* (by C. W. Cotman and J. L. McGaugh) and *Physiological Psychology* (by T. S. Brown and P. M. Wallace), are available in the UK in a cheaper international edition. Both are published by Academic Press, and the prices of the international editions are £6.25 and £6.40, respectively.

Richard of York's battles not in vain . . .

Ann Henderson-Sellers

Rainbows, Halos, and Glories. By Robert Greenler. Pp.240. ISBN 0-521-23605-3. (Cambridge University Press: 1981.) £15, \$24.95.

ACCURATE recollection of the ordered colours of the rainbow by means of the nursery sentence, "Richard Of York Gave Battle In Vain", is a necessary but insufficient basis for understanding in the eyes of Robert Greenler. He wants us also to be moved to wonder at the mysteries of optical phenomena in the sky. His book is reminiscent of a fascinating, extended tutorial with a professor of physics who, fortunately, combines hard scientific research with an almost childlike delight in beautiful things. During his discourse we are prompted to exclaim at magnificence, explain apparent enigmas and reinterpret historical events through computer simulations of atmospheric optics.

The book divides into three sections dealing with rainbows and ice-crystal optics, and then simulation and understanding of historical and modern sky displays, before finally returning to colourful mysteries, myths and mirages. Throughout, facts and frustrations are, like the individual rainbow we each view, enriched by Greenler's personal experiences. For instance, he illustrates the concept of the constant angular size of rainbows by a description of his doomed attempts to photograph the, apparently, miniature bow formed in the spray of his garden hose. The author is clearly "hooked" on his subject; he admits to seeing and searching for rainbows everywhere and has even captured infrared bows on film.

Despite his obvious love of rainbows it is with the optical displays formed by ice-crystals that Greenler is most strongly associated. As a physicist who has spent extended periods in Antarctica and the Arctic his fascination with this topic is understandable, but it is frustrating to realize that many of the sky effects he illustrates in this section are seen only rarely in more populated regions. This is because the huge variety of optical effects discussed depend upon the presence of simple hexagonal ice-crystals while most of us see our ice-crystals as snowflakes. This pinpoints the only real criticism that I can make of the book: it is not focused on a specific readership. As a meteorologist I am frustrated by the fact that the weather which, after all, is creating most of these effects, is mostly ignored. This is a pity because it makes an otherwise outstanding book tricky to use directly in taught courses. Despite the fact that the book is not intended as a textbook, with little additional effort topics could have been related to other disciplines. For instance, the very thin cirrostratus clouds

responsible for the 22° halo, itself generally considered as a forerunner of stormy weather, could also be considered to be a sky phenomenon. I can testify to having used halo occurrence to convince a non-meteorologist that a blue sky could be totally cloud-covered. Despite these reservations I was drawn on by the enthusiasm of the writing. Indeed, I spent some time testing the author's assertion that hexagonal ice-crystals have preferred orientations in the atmosphere, to the extent of spinning and dropping computer cards around my office.

With a flair for detective work which Holmes would have envied, Greenler considers his computer simulation of the famous "Arcs of Lowitz". These composed part of the historic sky display seen and sketched by Tobias Lowitz in St Petersburg in 1790. The accuracy of the simulated curves seems remarkable but Greenler remained dissatisfied. By returning to the written description he has demonstrated that the sketch is not a "snapshot" but a composition of a number of hours of viewing. This revelation solves the outstanding problems

which depended upon knowledge of the solar zenith angle.

The final 50 pages are devoted to colourful optical phenomena in both senses of the word. We are treated to pictures, and explanations, of white and black clouds in the same scene and then yellow, green, blue and red clouds. Such comparatively "every-day" sights pale before the "Specter of the Brocken". This apparition takes the form of a misty giant around whose head are a series of brilliant halos. Many air travellers have seen the same effect, a glory, which simply requires the observer to be in direct sunlight with his anti-solar point projected onto a cloud.

In the opening pages of the book the author states, "I have chosen, knowingly to include myself . . . my delight at seeing . . . my excitement in the process of discovery". Despite my slight unease at such an apparently selfish exercise I shall use and recommend the text. Like the Pied Piper of Hamelin in his rainbow-coloured coat, Greenler charmed me into following and will, I am sure, similarly entrap other readers. It appears that the white rose of the house of York was not lost in vain. □

Ann Henderson-Sellers is a Lecturer in Meteorology at the University of Liverpool, currently studying global climate at the Goddard Institute for Space Studies, New York.

Evidence of intent

John Stachel

The Greatest Power on Earth: The Story of Nuclear Fission. By Ronald W. Clark. Pp.342. ISBN 0-283-98715-4. (Sidgwick & Jackson: 1980.) £8.95. To be published in the USA by Harper & Row, \$13.95.

SHORTLY after the accident at Three Mile Island, a Japanese journalist visited a centre for evacuees from the vicinity of the reactor. As startling as the accident itself was the fact that a reference to Hiroshima drew a blank from those he interviewed; none of them recognized the name. Such ignorance in itself justifies more books on the story of the atomic bomb, which is basically the topic of Clark's book.

Over two-thirds is devoted to the story, from the discovery of atomic fission in late 1938, through the wartime efforts to develop nuclear weapons in Britain, the United States, Germany and Russia, up to the explosion of the first Russian bomb in 1949. A 40-page prologue to the book attempts a potted history of nuclear physics; while the last 40 pages wind up the story with brief accounts of the development of the fusion or hydrogen bomb, the attempts to use fission as an energy source and the growing awareness of radiation perils associated with peaceful and military uses of nuclear power.

The prologue is not too successful. It betrays a shaky acquaintance with the

scientific literature, resulting in numerous small errors and confusions. The placement of emphasis, too, is sometimes curious.

The core of the book demonstrates a wide acquaintance with the atomic bomb literature, both original documents (including some discussed in print for the first time) and later memoirs and analyses. On the whole, it seems a thorough and reliable account, correctly giving greater emphasis to work done in Britain than many American sources. This portion of the book commends itself to those in search of a readable, fairly non-technical account of the scientific, diplomatic and military events leading up to the "Birth of the Bomb" (to use the title of Clark's previous account of Britain's underrated contribution to the early phases of the story). Those in search of a deeper analysis of the potential background to the decision to drop the two bombs on Japan are still advised to consult Martin Sherwin's *A World Destroyed* (Knopf, 1975).

The frequent use of hyperbole is annoying: "In 1905 Einstein erupted into a surprised world" (elsewhere "an astounded scientific world" receives Einstein's 1905 relativity paper); Russian nuclear research is interrupted in 1941 because "German tanks with muffled treads moved through the darkness across the Russian frontier", and so on. There are also a few curious phrases: "The predominance of Jewish nuclear physicists on the Continent — in strong contrast to the situation in Britain —

has long been a subject for discussion"; by whom one wonders?

But the more serious problems of the book begin with its title. It is basically the story of nuclear fission, as its subtitle indicates, yet the title clearly refers to nuclear fusion — a reference reinforced by the exploding hydrogen bomb on the jacket cover. Careful reading of the text makes clear the distinction between fission and fusion, but hasty readers may be left confused.

More important is the question of the distribution of emphasis inherent in the structure of the book. Surely, the critical element in our current nuclear predicament is the combination of the fusion or hydrogen bomb, with its capacity for unlimited growth in explosive power, and a missile technology capable of accurately placing such bombs on target within minutes of launching. But no sense of the crucial nature of this development emerges from the brief concluding section of the book. Indeed, what is the moral of Clark's story? He says all the "right" things about the perils of nuclear weapons, and dangers of nuclear radiation — except for one sentence in his concluding paragraph, which plays into the hands of those currently trying to prepare the public for "limited" nuclear wars. He says,

It would be tragic if a limited nuclear exchange, halted after each side had experienced the agony of one city put to the bomb, were necessary to bring men to their senses and to realize, at last, that there are no victors in nuclear war.

Not only would it be tragic (reminding one of the man being led to his execution who assures the audience that this will certainly teach him a lesson!); it is so unrealistic a prospect, as Lord Mountbatten came to realize, that its contemplation as a pedagogical device seems irresponsible.

Finally, there is the broader question raised by the context in which Clark's account places the story of nuclear weapons. Focusing attention in the opening sections exclusively on the development of nuclear physics fosters the all too prevalent view of science and technology as independently evolving structures. For more direct evidence of this outlook, one need look no further than the opening sentence of Lord Zuckerman's preface:

The keystones of our century's history are those scientific and technological advances — radio, the internal combustion engine, antibiotics — which have determined the pattern of life which half the world now leads, and for which the other half strives . . .

or the opening paragraphs of Clark's first chapter. After evoking the political and social turmoil of late 1919, inflation and starvation in Central Europe, civil war in Russia, soldier's soviets in Britain, American rejection of the League, Clark continues:

Yet this record of battery and barricade, which added a bloody touch to the first autumn of

peace, was of infinitely less importance to the future of the human race than three events in the worlds of astronomy and physics.

More disturbing than Clark's particular choice of three events (about which more could be said) is the very idea of singling out some group of scientific and technological developments and setting them in apposition to the symptoms of vast social, economic and political upheavals in evaluating their "importance to the future of the human race". I think the evidence exists, in Clark's book and elsewhere, to sustain Raymond Williams' recent judgement:

In the case of nuclear weapons, nothing is more evident than that they were consciously sought and developed, and have continued to be consciously sought and developed. It is true that, as so often in modern technological innovations, much of the basic research had been done for quite other reasons, without foreseeing this particular result. But again as in many other comparable cases, the crucial moment of passage from scientific knowledge to technical invention, and then from technical invention to a systematic technology, depended on conscious selection and investment by an existing social order, for known and foreseen purposes (*New Left Review* 124, Nov-Dec. 1980, p.365).

Until we grasp this element of *intention* in the process of research and development in modern technology, as Williams puts it elsewhere, I fear that efforts to organize movements to control the uses of that technology will be hampered by the feeling that such developments really are inevitable. And in the case of nuclear weapons, I am not certain how much time we have left to learn the lesson. □

John Stachel is Professor of Physics at Boston University, currently editing the *Einstein papers* at the Institute for Advanced Study, Princeton.

More pie in the sky

Paul Davies

Gravity, Black Holes, and the Universe. By Iain Nicolson. Pp.264. ISBN 0-7153-7849-X/0-470-027111-6. (David & Charles/Halsted: 1981.) £10.95, \$24.95.

BOOKS on black holes and the Universe range from postdoctoral research treatises to the wildest distortions of science fiction and the occult. Since the explosion of interest in astronomy in general and black holes in particular, which began during the 1960s, the market has been flooded with literature on matters cosmic. How can the discerning reader tell where to go for both reliable and entertaining treatment?

Having spent many years myself on cosmological and astrophysical research, I do not envy the task of the professional writer who has to attempt to master the technical literature of a difficult and developing subject, and distil out the

features that the scientific community considers to be well established. Still more demanding is the task of grading the various conjectural topics, some of which can be bizarre, into degrees of seriousness.

Take one example: travel through a black hole. Certain solutions of Einstein's field equations display the weird property that they can be continued into regions of space and time that are forever inaccessible to us unless we fall into a black hole. These "other" spacetimes (infinite in number) are sometimes called other universes. Do they exist?

In the sense that certain idealized model spacetimes admit other universes, then the subject is a serious one, and schematic representations of these universes appear in textbooks and mathematical monographs. In spite of this, the models concerned are not supposed to apply to the real world, because they involve certain simplifications (for example, the absence of matter) that manifestly fail in reality. The prevailing opinion is that a more realistic model would not display these other universes, and that an observer who falls into a black hole will not, after all, emerge in a hitherto unknown spacetime, but will crash to oblivion in a region of unlimited tidal stress.

A writer may cover his subject thoroughly yet still miss such subtleties and proclaim that modern physics definitely predicts the existence of other universes. Certainly a glance at a textbook could give that impression. Iain Nicolson is to be congratulated that on this, and many other contentious and delicate issues, he has not only got the facts right, but has conveyed the correct degree of scepticism or reliability appropriate to each conjectural topic.

There is much more to this book than black holes, however. The author takes us gently through Kepler's laws, newtonian mechanics and special relativity, before getting on to the equivalence principle, curved spacetime and the general theory of relativity. He also applies his thoughts to the expanding Universe, and the intriguing topic of its ultimate fate.

The book is nicely illustrated with line drawings and photographs, and has a short bibliography. Although certain advanced topics, such as quantum aspects of black holes, are courageously included (and skilfully handled), the book should still appeal to general readers having only a nodding acquaintance with physics and astronomy.

My initial reaction to this book was to wonder where yet another treatment of this trendy subject could possibly fit in to an already crowded market. The answer is that it is one of the few books that maintains a reliable scientific perspective and a careful treatment, yet covers most of the "fun" topics in detail. □

Paul Davies is Professor of Theoretical Physics at the University of Newcastle upon Tyne.

The diversity of physics revealed

R.H. Dalitz

Encyclopedia of Physics. Edited by Rita G. Lerner and George L. Trigg. Pp.1,157. ISBN 0-201-04313-0. (Addison-Wesley: 1981.) \$99.50, £49.50.

THE publication of a new encyclopaedia of physics is an important event. The subject is growing rapidly, not only in accuracy of measurement and detail of understanding but also in scope. A well organized volume can quickly bring the reader up to date about developments in areas outside his own speciality where new viewpoints and techniques may have become established, and, overall, can significantly speed up the progress of physics itself.

This volume contains about 450 entries, some long (18 pages on symbols, units and nomenclature; 11 on elementary particles) and some short (a few less than a column). They range from instrumentation (particle accelerators, photomultipliers, transistors, thermocouples and so on), through classical physics (analytical dynamics, thermodynamics, acoustics, geomagnetism etc.), to quantum physics (liquid helium, atomic spectroscopy, nuclear structure, Josephson effect and so on). The articles are all by well-known physicists. They show great diversity in both objective and style, and there is considerable overlap between them.

How well will this new encyclopaedia meet our needs? Many of the major entries are really excellent — for example those on elementary particles, liquid crystals, vision and colour, and non-equilibrium thermodynamics, to mention just several. For greatest effectiveness, each major article should really provide a substantial reference list, but not all do. That on nuclear reactions, two pages long, provides an excellent classification of work in this subject but gives only one reference, whereas other articles, such as those on electronics, the Hubble effect, and relativity (special theory), give a wide range of graded references; perhaps this is overdone for solar energy, with one and a half pages of references for a six-page entry.

With a variety of viewpoints of the contributors, the multiple coverage could be a strength, but some consolidation would have obvious advantages. To give one illustration, the entries for baryons, hadrons, hyperons, and mesons, each about a page in length, could usefully have been combined into a four page article on hadrons, the other three entries each being reduced to one line, cross-referenced to hadrons. In addition, the coverage of different topics is not always in proportion to their relative importance. For example, the rather detailed contribution on weak neutral currents is given five pages, without any corresponding counterpart for weak charged currents. Also turbulence, one of the least understood areas of physics,

receives only one and a half pages.

Many of the articles read strangely for a newly published book, as if one is recalling a past time, some four or five years ago. For instance, quantum chromodynamics, our deepest and most promising theory of elementary particles, receives only one paragraph in each of the contributions on elementary particles, quantum electrodynamics, and unified field theory, despite the fact that it has provided an underlying basis for most of the theoretical work in elementary particle and high-energy physics over the past three years. The editors of an encyclopaedia face great difficulties, of course, but in this case they might have noted that a few of the articles called for some revision at the last moment.

The operational test of a reference book of this sort is whether information can be located rapidly. For this, the general index serves fairly well. I found that Higgs mechanism, radiocarbon dating, and

Peierls stress were in the index, as was Umklapp process, although with reference to p.732 where only articles on partial waves and partons are to be found. There is no entry for current algebra but the article on currents in particle theory is indexed nearby and led me to my goal. However there is no entry for Thomas precession, nor for photon multiplier, nor even for the Wigner-Eisenbud theory of the R-matrix and its application to resonance processes.

This encyclopaedia has a great deal to offer to scientists and engineers in general, as well as physicists. The underlying plan is sound and the articles generally authoritative. With so many first-rate contributions, it should have an assured place in all science libraries. It could be substantially improved, in my opinion, but, as things stand, it is the best compilation of its type available for physics today. □

R. H. Dalitz is Royal Society Research Professor in Theoretical Physics at Oxford University.

This is psychology?

John C. Marshall

Mindwatching. By H. and M. Eysenck. Pp.224. ISBN 0-7181-1937-1. (Michael Joseph: 1981.) £10.50. To be published in the USA by Doubleday.

THERE'S an old joke about how to select candidates for admission to psychology courses. The hapless individuals are asked the simple question "Why do you want to study psychology?" Those who reply "Because I want to *understand* people" are advised to read history or English literature; those who reply "Because I want to *help* people" are told to train as surgeons, farmers, garbage collectors or tax lawyers. Any other response secures admission to the psychology department.

Hans and Michael Eysenck would, I presume, disapprove. For them, the goal of professional psychologists is "to understand why people behave as they do, so that it will be possible to predict and change their behaviour"; this goal is to be achieved by controlled experimentation, the results of which are sometimes counterintuitive and thus undermine the sceptic's claim that my grandmother knew as much "psychology" as it is possible (or wise) for any mere mortal to comprehend. The Eysencks have accordingly produced this coffee-table volume in which the purported achievements, theoretical and practical, of experimental psychology are laid out for interested bystanders. Two questions therefore arise: First, to what extent are we presented with a fair summary of modern psychology? Second, are the social implications of current work outlined in a

responsible fashion?

The authors employ the stylistic device of describing in each chapter a so-called "key experiment", supplemented by anecdote and speculation. This is not in itself a bad strategy but it does lead the Eysencks into letting vaudeville prevail over rationality. For example, a section on "Attractiveness and jury decision-making" is introduced by the story of a famous nineteenth-century trial in which a woman accused of poisoning her lover was somewhat lucky to obtain a verdict of "not proven". The Eysencks speculate that since the defendant was "young, vivacious, and beautiful", these attributes may have influenced the jury and "helped to save her from the gallows". An experiment is then described in which mock juries who consider hypothetical cases of a relatively minor nature do indeed seem to be swayed a little by the physical attractiveness of the mock defendants. But the section then concludes with the claim that "most of the available research suggests that physical attractiveness has little or no effect on juries when serious crimes are involved". I can only make sense of this information by assuming the Eysencks believe that murdering one's lover is not a serious crime.

On other occasions the authors are unable to resist the temptation of a spectacular but fatally flawed study. They thus feature Rosenham's report in which a group of normal people were admitted to mental hospitals after falsely complaining that they heard voices and asking to be admitted. Following an average stay of 19

days the pseudo-patients were discharged with a diagnosis of "schizophrenia in remission". On the basis of this and a number of related pranks, Rosenham concluded that "It is clear that we cannot distinguish the sane from the insane in psychiatric hospitals". Very amusing, no doubt, but normal people who are not in the pay of psychologists either don't hear voices or, if they do, are not sufficiently worried by it to seek admission to hospital. The Eysencks themselves quote Seymour Kety's withering comment — "If I were to drink a quart of blood and, concealing what I had done, come to the emergency room of any hospital vomiting blood, the behaviour of the staff would be quite predictable. If they labelled and treated me as having a bleeding peptic ulcer, I doubt that I could argue convincingly that medical science does not know how to diagnose that condition" — but they are not thereby inhibited from describing Rosenham's jest as a "key experiment".

Not all of the selections are as silly as Rosenham's study. The book contains quite sensible sections on eye-witness testimony, divided attention and forgetting. Similarly, the chapter on attempts to teach sign language to chimpanzees is an improvement on many popular treatments of the topic. But these islands of sanity are lost in the surrounding sludge of trivia and titillation. What can one say of a book which proclaims that psychology should "be regarded as the most important of all the sciences" and then reports that "Small breasts were preferred by men with fundamentalist religious beliefs suffering from mild depression"?

There is little improvement when the Eysencks draw out the practical implications of their choice of topics. We are informed that a process of conditioning can turn our children into "well-integrated, honest members of society" and referred to Chapter 6 for the experimental evidence that this is so. Chapter 6, however, reports that obsessive and compulsive disorders can be treated by "behaviour therapy". In Chapter 22, a discussion of Hans Eysenck's "theory" of personality is held to support the claim that "The egalitarian principle underlying socialism is in conflict with biological reality and can only be enforced by dictatorship". Exactly how this conclusion is reached remains shrouded in mystery. An appallingly ill-controlled "experiment" on the effects of two types of play-group prefaces the remark that "the advocates of progressive teaching . . . have turned our pre-school groups, kindergartens and primary schools into permissive places of no work and all play, from which all socializing and restraining influences have been banned".

One can only hope that this grubby little opus will sink without trace. It totally fails to represent those aspects of psychology where results of some intellectual depth have been obtained in the study of perception, language and skilled action; it similarly ignores many areas of occupational, clinical and educational research where careful, thoughtful work has served to enhance the quality of life. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford.

Check-list of Birds of the World. For example, Terres leaves the titmouse family intact and includes the long-tailed tits (Aegithalinae) and penduline tits (Remizinae) as subfamilies rather than treating them as distinct families as has been recently practised. On the other hand, the thrushes are retained in their own family, Turdidae, and there is no mention of their current placement as a subfamily in the Old World flycatcher family, Muscicapidae.

While the book has many virtues it falls short of the stated goal of providing information on all the birds that have nested or been sighted in Alaska, Canada, Greenland, Bermuda, Baja California and the 48 contiguous states of the USA. Omissions are conspicuous, for example, in the Alaskan avifauna. While the 1978 update of the status of Alaskan birds is referenced in numerous places, and 1980 Alaskan references are also given, it is annoying that no less than 22 Alaskan species given in the 1978 publication are ignored. For instance, while the willow warbler (an Old World warbler) is discussed, nothing is said about its North American status. Only its Eurasian distribution is mentioned, so readers are not aware that it has occurred as an accidental in Alaska.

Another flaw is that the book's attractiveness is compromised by lumping colour plates together in sections. For example, located in the middle of the thrush family species accounts are 31 pages of colour plates of tanagers, thrush family, titmouse family, trogon, tropicbirds, troupial (American blackbird) family, turkey, vireo family and New World vultures and condors — rather a potpourri even though they are in alphabetical order. As with most undertakings of this magnitude other errors, mainly of editorial style, creep in. Under "Fecal Sac" we are referred to "Nests and Nestlings" but there is nothing on the sac there. A second reference under "Fecal Sac" is "Nest Sanitation" which is there, but under that title we are referred further to "Young and Their Care" which contains a discussion of nest sanitation and fecal sacs. Further minor inconsistencies are the listing of "Courtship" with sub-topics such as *courtship rituals* but not *courtship feeding*. This latter appears as a completely separate topic followed by "Courtship Flight", which then only refers the reader back again to "Courtship" and "Flight".

Despite my less-than-total enthusiasm with the *Encyclopedia*, on balance Terres has done a remarkable job in producing a scientifically accurate tome relatively free of typographical errors. With over 7 million bird-watchers in the US alone (as of 1970), of which at least 3 million were considered to be seriously interested in field ornithology, there should be a receptive audience to savour this book. □

Clayton M. White is a Professor of Zoology at Brigham Young University, Provo, Utah.

Find the (North American) bird

Clayton M. White

The Audubon Society Encyclopedia of North American Birds. By John K. Terres. Pp. 1,109. ISBN 0-3944-6651-9. (Knopf: 1980.) \$60 US only.

IN 1896 Professor Alfred Newton published his *Dictionary of Birds*. Thus a mould was cast. But it took another 68 years, until 1964, through the British Ornithologists' Union and under the editorship of A. Landsborough Thomson, for another such work to appear: *A New Dictionary of Birds*. Now, only 16 years later, another giant encyclopaedia is on the market. To me, this current volume appears to be an attempt to update a considerable body of material from the now elusive *New Dictionary*. Its scope, however, is mainly restricted to North America.

The book is designed to be a single-volume work for both layman and professional, but I suspect the lay will benefit most. There are nearly 6,000 alphabetically arranged entries and cross references, illustrated with 875 full-colour

photographs of bird species and over 800 black-and-white drawings. The scope of the entries is impressive: Birdhouse, Bird Lice, Birds as Collectors, Birds as Watchdogs, Bird Skin, Bird-Watcher, Bisexuality, Bittern, Blackburne, among the "B's" for example. Abbreviated life histories of 847 birds are provided. For each species are given a description of the Latin or Greek origin of the scientific name (a most valuable feature), the bird's geographical range, colour and distinctive morphological features, habitat, feeding habits, other names, nest, eggs, and other features of interest (such as who parasitizes whom). Terres cites over 4,000 references, ranging in date from the turn of the 1800s to 1980, all of which are listed in a bibliography. The bulk of several alphabetical sections is taken up by species accounts. The taxonomic treatment is fairly current by both European and American standards, at least in common names, but is more conservative and departs somewhat from the nomenclature of Peter's

REVIEW ARTICLE

Localization and electronic properties in amorphous semiconductors

J. Mort

Xerox Webster Research Center, Webster, New York 14580, USA

J. Knights

Xerox Palo Alto Research Center, Palo Alto, California 94304, USA

The transport and photogeneration of different classes of amorphous semiconductors are considered. These properties are affected by localization associated with long-range disorder and defects which in turn reflect the underlying chemical nature of each class.

PRESENT interest in the electronic properties of amorphous semiconductors is the result of several related factors. The demonstration^{1,2} that amorphous Si, (a-Si), like the crystalline form, can be doped to form either a *n* or *p* extrinsic semiconductor suggested new ways of using amorphous materials as electronic materials.³⁻⁶ Other classes of amorphous solids than this tetrahedral material, have been extensively studied in recent years, including chalcogenides such as amorphous selenium (a-Se), amorphous arsenic triselenide, (a-As₂Se₃) and amorphous molecular or organic solids⁷⁻¹⁰. It is now clear that although amorphous solids have certain characteristic features in common, there are also distinctive differences. Our purpose here is to compare and contrast the electronic properties of these three groups of amorphous solids—the chalcogenides, the tetrahedrally bonded materials and molecular systems.

The most obvious consequence of disorder in amorphous solids is that intrinsic localized states exist within the energy gap. The experimental and theoretical problems of understanding the effect of disorder on the photoelectronic properties of solids are considerable, but much progress has been made in the past decade in three directions. Light has been thrown on the origin and nature of localized states, and their relationship with the specific chemical nature and the bonding of the constituent building-blocks. The mechanism of charge transport and photogeneration has been partially understood; quantitative studies of the consequences of doping have begun. In what follows, these three areas will be discussed for the three general classes of amorphous solids.

Nature of localized states

In the development of crystalline semiconductors, major conceptual advances such as the understanding of band-gaps arose from the recognition of the long-range periodicity and symmetrical arrangement of the atomic building blocks—intrinsic features which are essentially independent of the method of preparation. Most technical applications of crystalline semiconductors and their associated physics are, however, connected with deviations from idealized order. The localized states within the band-gap produced by defects, impurities, surfaces and interfaces control many properties, particularly those involving carrier transport and generation-recombination processes; such perturbations are strongly influenced by both the method of preparation and subsequent treatment.

Viewed in this way amorphous semiconductors exhibit remarkably similar features. The existence of band gaps in the absence of long-range periodicity has a firm theoretical basis in tight-binding approaches. It has been predicted^{11,12}, based on Anderson's work on localization¹³, that even in the absence of defects the potentials fluctuating because of the disorder would result in localized tail states at the conduction and valence band edges. Figure 1 shows that the density of states distribution in an amorphous solid can be divided into three regions. For energies greater than ϵ_c and less than ϵ_v , electronic states are extended or band-like in character. For energies close to these values, the scattering length of the electron wave function approaches interatomic distances. Conduction then occurs as a Brownian diffusive motion with estimated mobilities between 1 and

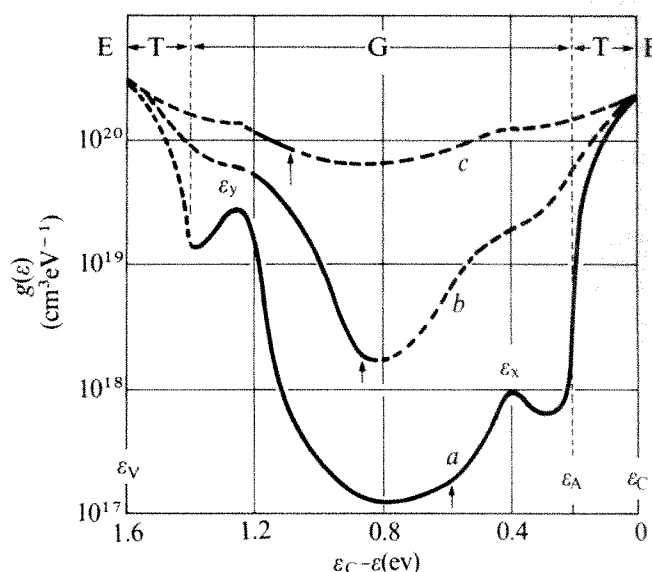


Fig. 1 Density-of-state distributions for a-Si specimens. *a*, Glow discharge specimen, $T_d \approx 520$ K; *b*, glow discharge specimen, $T_d \approx 350$ K; *c*, evaporated specimen. The solid lines indicate results obtained from field-effect experiments and the arrow on each curve shows the position of the Fermi level. E, extended states; T, tail states; G, gap states. (After ref. 15.)

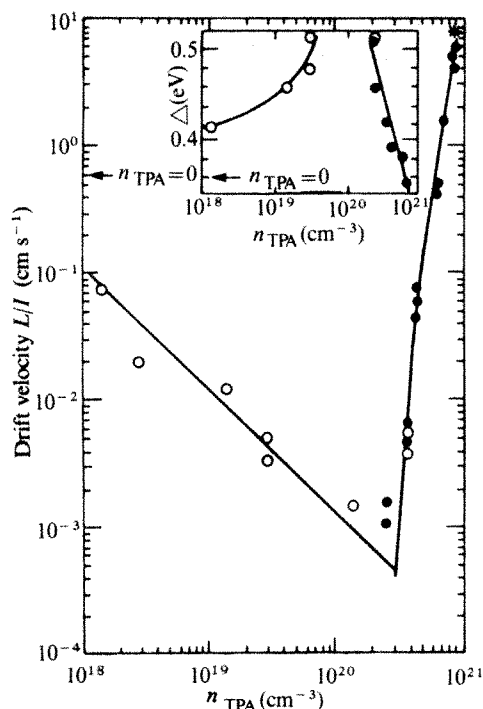


Fig. 2 Concentration dependence of the drift velocity of holes in a mixed doped polymer showing the trap-controlled hopping process. NIPC concentrations: $\circ$, $1 \times 10^{21} \text{ cm}^{-3}$; $\bullet$, 0; $*$, $2 \times 10^{20} \text{ cm}^{-3}$. The points correspond to samples with varying concentrations of TPA, which has a lower ionization potential than NIPC. (After ref. 25.)

$10 \text{ cm}^2 \text{ V s}^{-1}$ (ref. 14). Just below ϵ_c and above ϵ_v there exist the tail states due to Anderson localization that are associated with satisfied but distorted bonds. Anderson localization begins when U_0/J (where U_0 is the range of site energy disorder and J is the transfer integral between nearest neighbours) exceeds a critical value at which an electron at a given site at temperature $T = 0$ becomes localized, that is, it can no longer diffuse away. Defects arising from incomplete or dangling bonds or impurities give rise to the deep lying gap states. As in the crystalline case, the nature and density of these states depend on conditions of preparation and treatment.

The determination of detailed density of state distributions has proved difficult. In the case of amorphous Si some information has come from studies of the field-effect¹⁵ as shown in Fig. 1. Although quantitative details and the relative contributions of surface and bulk states are uncertain, the general features are consistent with independent studies of conductivity, thermopower and drift mobility. The distribution and number of localized states in a-Si:H prepared by glow discharge of silane may be controlled by the preparation conditions (Fig. 1). This is partly because hydrogen satisfies dangling silicon bonds at defects in the extended amorphous network. Many of the associated electronic defect states are thereby removed from the forbidden gap, which suggested that impurity doping above the residual levels might be possible^{1,2}.

These features are general and may be applied to the covalently bonded chalcogenides. However, all efforts to observe a significant field-effect have failed¹⁶. This is surprising¹⁷ because the efficiency of amorphous Se in electrophotography indicates a substantial carrier range of $\sim 50 \mu\text{m}$, which implies gap state densities $\sim 10^{14} \text{ cm}^{-3}$. Recent intrinsic defect models for gap states, which take account of the specific atomic bonding of chalcogen atoms coupled with the law of mass action, may resolve this conflict^{18,19}. The valence alternation pair defects (VAPS) lead to a thermodynamic equilibrium of D^+ and D^- states, which effectively make the Fermi level independent of

the actual density of the state; in other words the Fermi level cannot be moved relative to the bands either by an applied field as in the field effect experiment or doping.

In disordered molecular solids, the building blocks are molecules. Even in molecular crystals the weak Van der Waal's forces between molecules lead at best to very narrow energy bands. In this case, the disorder, associated with pendant group polymers such as poly(*N*-vinylcarbazole) or molecularly-doped polymers means that extended states cannot be involved⁸. Such systems, particularly molecularly-doped polymers, therefore must consist of purely localized states which are associated with the dopant molecules⁸, in this way they are almost ideal for studying phenomena associated with localized states.

Electronic transport

The most unambiguous information on electronic transport in amorphous materials has come from measurements of the drift mobility by the time-of-flight technique. In this method the time, t_T , for a sharply defined spatially and temporally photoexcited charge sheet to cross a sample of known thickness L , with an applied bias field E , is measured. Because $t_T = L/\mu E$, the carrier mobility μ can be directly determined. Table 1 gives typical values of the drift mobilities for both electrons and holes in representative amorphous materials. In all cases the mobilities are unity or less and are thermally activated. There are various explanations for the nature of this electronic transport. No conduction can occur in the localized tail and gap states at $T = 0$. At finite temperatures, however, thermally activated hopping transport between localized states can occur and the mobility $\mu(\epsilon)$ at an energy ϵ can be written²

$$\mu(\epsilon) = \frac{1}{6} \frac{eR^2(\epsilon)}{kT} \nu_{ph} \exp(-2\alpha R) \exp(-W/kT) \quad (1)$$

where R is the average distance between sites at energy ϵ and is determined by the density-of-state distribution. The term $\exp(-2\alpha R)$ describes the overlap of the wave functions on neighbouring hopping sites and α^{-1} is a measure of the spatial decay of the wave function^{2,7}; $\nu_{ph} \exp(-W/kT)$ represents the probability that the localized carrier hops to a site with energy higher by W . This transport mechanism solely controlled by hopping from one localized state to another should be observed in amorphous materials. However, before doping studies proved possible, this mechanism could only be deduced with some ambiguity from the combined observation of low mobility values and thermal activation as these features are also characteristic of transport involving extended state transport modulated by trapping in energetically shallow traps²⁰.

Direct observation of pure hopping between localized states became possible with the advent of molecularly-doped polymers²¹. These are otherwise electrically-inert polymers into which molecules can be dispersed at high controllable concentrations $< 10^{21} \text{ cm}^{-3}$ and constitute a random ensemble of molecules which have essentially no mutual interaction. As a result they consist of pure localized states without the complication (and advantage) of extended states. Measurements of the effective mobility by the time-of-flight technique as a function of

Table 1 Representative hole (electron) drift mobilities at room temperature for examples of different types of amorphous materials

	$(\text{cm}^2 \text{ V}^{-1} \text{ s}^{-1})$	$(\text{cm}^2 \text{ V}^{-1} \text{ s}^{-1})$
a:Se	$(5-8) \times 10^{-3}$	$0.11-0.19$
a:As ₂ Se ₃	—	$\sim 10^{-5}$
a-Si:H	0.1	5×10^{-4}
PVK	—	$< 10^{-6}$

See refs 2, 8, 27.

Table 2 Amorphous materials are different and may be categorized according to how the conductivity ($\sigma = ne\mu$) is affected by doping

Chalcogenides	Tetrahedral	Molecularly doped polymers
σ , relatively weak dependence	σ strong dependence	σ strong dependence
$\Delta\sigma$ related to $\Delta\mu$	$\Delta\sigma$ related to Δn	$\Delta\sigma$ related to Δn and $\Delta\mu$
(Fermi level immobile)	(Fermi level moves)	

An example of such dispersive transits in a-As₂Se₃ are shown in Fig. 3. In Fig. 3a a transit signal, showing the photoinduced transient current versus time in linear scales, is essentially featureless. The same transit signal in Fig. 3b is displayed with logarithmic current and time scales. The predicted change in slope when the leading carriers have transited the sample is indicated at t_T which shifts to shorter times at higher fields, E . The concept of mobility now loses its meaning as no average velocity of the charge packet is definable, and the predicted thickness-dependent mobilities have been observed. This type of dispersive phenomena appears in recombination processes, time dependent luminescent decay and trapping²⁹.

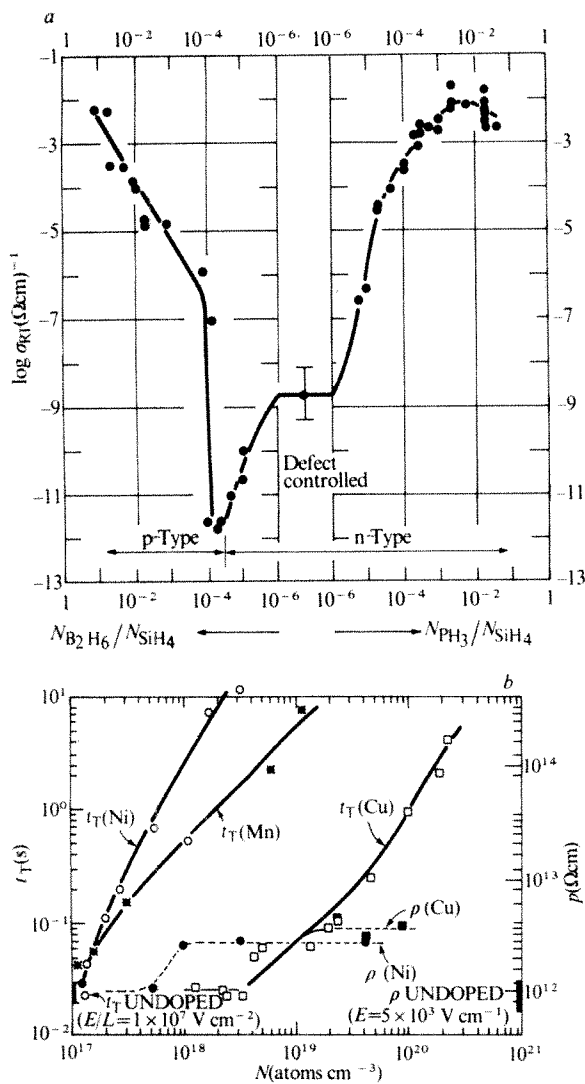


Fig. 5 a, Room-temperature conductivity σ_{RT} of n and p-type amorphous Si specimens as a function of the gaseous impurity ratio. For the right-hand curve, this is the number of phosphine to silane molecules in the gas mixture used for specimen preparation. On the left it is the corresponding diborane to silane ratio. The centre refers to undoped specimens. (After ref. 1.) b, log-log plot of transit time and resistivity ρ versus concentration N of: $\circ$, $\bullet$, Ni; $\ast$, Mn; $\square$, $\blacksquare$, Cu doped a-As₂Se₃. (After ref. 37.)

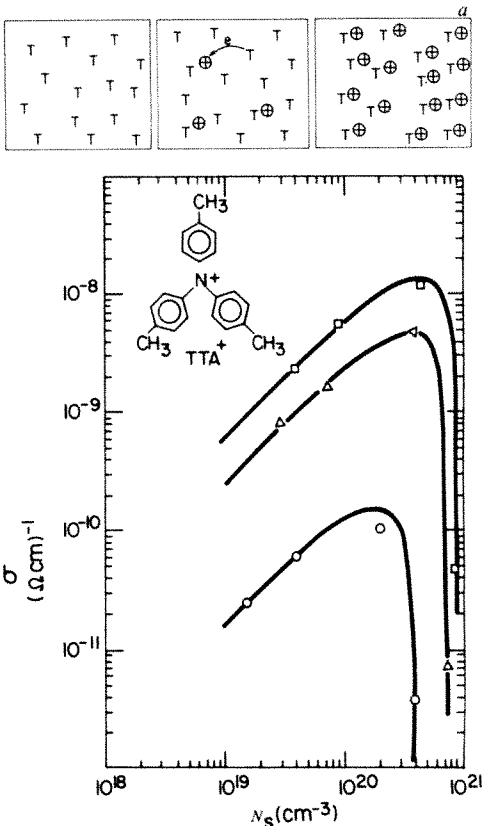


Fig. 6 a, Chemical control of conductivity in a molecularly-doped inert polymer matrix. Fractional oxidation of neutral transport molecules T is achieved by controlled addition of an oxidizing agent SbCl₅. The oxidized transport molecules are radical cations T⁺ which constitute free carriers. Progressive oxidation of the neutral molecules leads to increasing free carrier concentrations at the expense of the neutral molecules necessary for the transport mechanism. b, Dependence of conductivity on concentration of oxidized transport molecules N_s . The molecule tri-p-tylamine (TTA) is dispersed in Lexan[®] polycarbonate. The total molecule concentrations were: $\square$, $8.6 \times 10^{20} \text{ cm}^{-3}$; $\triangle$, $7.2 \times 10^{20} \text{ cm}^{-3}$; $\circ$, $0.86 \times 10^{20} \text{ cm}^{-3}$. (After ref. 40.)

Photogeneration

Photogeneration is a critical process in any photoelectronic material. The way in which absorbed photons produce photogenerated carriers is affected by disorder within the material. Geminate recombination is likely to be a limiting step in the photogeneration process in materials where the mean free paths of photoexcited carriers are characteristically low³⁰—a natural condition of amorphous materials. Figure 4a illustrates this photogeneration process whereby the initial photoexcitation leads to electron-hole pairs, which diffuse within their mutual Coulomb well. There is a finite probability, which decreases with applied field, that the pair will combine at the initial excitation site rather than dissociate into a free electron and a hole. Three critical parameters are; Φ_0 , the efficiency of production of thermalized electron-hole pairs per absorbed photon; r_0 , a characteristic thermalization length (related to the mean free path); and r_c , the coulomb radius or the distance at which the coulomb energy $\sim kT$. Precise measurements³⁰ for a-Se indicate Φ_0 (the maximum possible quantum efficiency) is unity and there is evidence that this also applies to a-Si (refs 31, 42) and a-As₂Se₃ (ref. 32). For disordered organic systems, however, Φ_0 can be ~ 0.1 or even smaller. This reflects the existence of intramolecular relaxation processes, radiative or non-radiative, which compete with the photogeneration or dissociation of the photoexcited molecule³³. Figure 4b shows a treatment of this problem due to Onsager³⁴ for a material with $\Phi_0 = 1$ and the dielectric constant of a-Se ($\epsilon = 6.0$). The various curves are for

different assumed values of r_0 . For a given value of r_0 the quantum efficiency remains field independent up to $\sim 10^4 \text{ V cm}^{-1}$ after which it becomes field dependent and approaches unity in a sufficiently high field. As r_0 increases the initial or zero field quantum efficiency increases and the subsequent field dependence decreases. Where a curve for a particular value of r_0 lies on the absolute quantum efficiency axis depends on the coulomb radius of the material which is determined by the dielectric constant. Typical coulomb radii are $\sim 49 \text{ \AA}$ for a-Si, $\sim 100 \text{ \AA}$ for a-Se and $\sim 250 \text{ \AA}$ for most organics, this would indicate descending order for the photogeneration efficiency for comparable values of r_0 and Φ_0 .

Experimental data for a-Se shown in Fig. 4b explain the macroscopic observation that a-Se is more panchromatic in absorption than in photosensitivity³⁰. This arises because r_0 being wavelength dependent ranges from $70 \text{ \AA}$ at a wavelength $4,000 \text{ \AA}$ to $10 \text{ \AA}$ at a wavelength of $6,000 \text{ \AA}$. Recent evidence shows that photogeneration in a-Si:H and a-As₂Se₃ may also be geminate^{31,32}. Unlike the case of a-Se, however, r_0 has at most a very weak wavelength dependence. A field-dependent quantum efficiency has implications for ultimate solar cell conversion efficiencies, this is because of the spatial field distribution within the barrier of Schottky or p-n cells which, with the absorption profile of the light, leads to photons being absorbed in a range of fields and effectively leads to a spatial distribution of quantum efficiencies³⁵.

Doping of a-semiconductors

One of the most important advances in the field of amorphous materials has been the demonstration that n or p type a-Si:H and a-Ge:H could be produced by appropriate doping; there is no evidence that this applies to chalcogenides, although it is possible in a-C:H (ref. 36). It seems possible to group amorphous materials according to the way in which doping affects their electrical conductivity σ , as indicated in Table 2. Detailed studies of doping in amorphous Si and, to a lesser degree, in a-carbon indicate that σ is strongly dependent on doping. Conductivity, thermopower and drift mobility measurements demonstrate that this is associated with a shift in the Fermi level which affects both the type and density of carriers controlling the conductivity (see Fig. 5a)². On the other hand, relatively small changes in the transport occur, except for special cases such as inter-donor hopping in compensated samples. By contrast, extensive doping studies of chalcogenides, for doping levels well below the alloy range, do not result in significant movement of the Fermi level or change in conductivity³⁷. However, substantial and systematic changes in the drift mobility of holes have been reported in a-As₂Se₃. Because the drift mobility can be the microscopic mobility modified by trapping, the fact that these changes in the drift mobility are not reflected in the conductivity (see Fig. 5b) indicates that the impurities are changing the density of traps which controls the mobility³⁷. This

insensitivity of the position of the Fermi level in chalcogenides to doping is consistent with the absence of a field effect and with other defect models^{18,19} based on valence alternation pair defects, D^+ and D^- , which are positively and negatively charged defects whose creation is governed by the law of mass action. Such defects produced under thermal equilibrium effectively pin the Fermi level. It is possible that doping of chalcogenides in non-equilibrium conditions could result in an unpinning of the Fermi level and production of extrinsic semiconductors. Whether this can be done in a stable manner is unclear.

Recent studies confirm that the densities of defects can be affected by heating³⁸. Thus photoelectronic properties of chalcogenides immediately after preparation are dependent on thermal history. Characteristic equilibrium behaviour, independent of the method of preparation is, however, achieved after annealing^{38,39}. The latter are consistent with Spear's observation that the most remarkable feature of chalcogenides is the reproducibility of their mobility values, a feature probably unmatched by even the best crystalline semiconductors¹⁶. This then, seems to be related to the ability of chalcogenide glasses to self-anneal defects efficiently.

In the disordered organic materials, doping is accomplished by the controlled chemical oxidation of some of the dopant molecules⁴⁰. These radical cations are the equivalent of 'free' carriers because electrons can hop from neighbouring neutral molecules and thereby the 'free' hole propagates (see Fig. 6a). Thus at low oxidation levels the conductivity rises as the 'free'-carrier concentration rises. However, because no extended states exist and free carriers are highly localized and associated with specific molecules, such free carriers are created at the expense of transport paths. In the extreme case, where the total oxidation has occurred, none of the neutral molecules necessary for the transport process remain and the conductivity must go through a maximum as indicated in Fig. 6b (ref. 40). It is possible to produce films containing $\sim 10^{20}$ cations cm^{-3} which form what Mott has termed a degenerate polaron gas⁴¹. In this case, therefore, the conductivity is controlled by changes in both free carrier concentrations and mobility where these are not independent.

Conclusions

We have outlined the transport and photogeneration properties in three general groups of amorphous materials which make up most of materials known as amorphous semiconductors. In particular the different effects of localization of electronic wavefunctions due to long range disorder and those due to defects have been considered. The picture that emerges is that although features characteristic of the disordered state do exist, amorphous materials exhibit a rich diversity of behaviour which reflects their underlying chemical composition. In this sense the similarities between disordered and crystalline solids are closer than was originally perceived.

1. Spear, W. E. LeComber, P. G. *Phil Mag.* **33**, 935 (1976).
2. Spear, W. E. *Adv. Phys.* **26**, 811 (1977).
3. Carlson, D. E. & Wronski, C. R. *Appl. Phys. Lett.* **28**, 671 (1976).
4. Spear, W. E., LeComber, P. G., Kinmond, S. & Brodsky, M. H. *Appl. Phys. Lett.* **28**, 105 (1976).
5. Gibson, R. A., Spear, W. E., LeComber, P. G. & Snell, A. J. *J. Non-Cryst. Solids* **35-36**, 725 (1980).
6. LeComber, P. G., Spear, W. E. & Garth, A. *Electron. Lett.* **15**, 179 (1979).
7. Pfister, G. *Phil. Mag.* **36**, 1147 (1977).
8. Mort, J. & Pfister, G. *Polymer-Plast. Technol.* **12**, 89 (1979).
9. Mort, J. *Adv. Phys.* **29**, 367 (1980).
10. Mort, J. *Science* **208**, 819 (1980).
11. Mott, N. F., *Electronic and Structural Properties of Amorphous Semiconductors* (eds LeComber, P. G. & Mort, J.) (Academic, London, 1973).
12. Cohen, M. H., Fritzsche, H. & Ovshinsky, S. R. *Phys. Rev. Lett.* **22**, 1065 (1969).
13. Anderson, P. W. *Phys. Rev.* **109**, 1492 (1958).
14. Cohen, M. H. *J. Non-Cryst. Solids* **4**, 391 (1970).
15. Madan, A., LeComber, P. G. & Spear, W. E. *J. Non-Cryst. Solids* **20**, 239 (1976).
16. Spear, W. E. *Amorphous and Liquid Semiconductors* (eds Stuke, J. & Brenig, W.) (Taylor and Francis, London, 1974).
17. Nielsen, P. *Phys. Rev.* **6**, 3739 (1972).
18. Mott, N. F., Davis, E. A. & Street, R. A. *Phil. Mag.* **32**, 961 (1975).
19. Kastner, M., Adler, D. & Fritzsche, H. *Phys. Rev. Lett.* **37**, 1504 (1978).
20. Dolezalek, F. *Photoconductivity and Related Phenomena* (eds Mort, J. & Pai, D.) (Elsevier, Amsterdam, 1975).
21. Mort, J., Pfister, G. & Grammatica, S. *Solid St. Commun.* **18**, 693 (1976).
22. Mort, J., Grammatica, S., Sandman, D. J. & Troup, A. J. *electron. Mater.* **9**, 411 (1980).
23. Allan, D., LeComber, P. G., & Spear, W. E. *Proc. 7th Int. Conf. on Amorphous and Liquid Semiconductors* (ed. Spear, W. E.) 323 (CICL, University of Edinburgh, 1977).
24. Emin, D. *Electronic and Structural Properties of Amorphous Semiconductors* (eds LeComber, P. G. & Mort, J.) (Academic, London, 1973).
25. Pfister, G., Mort, J. & Grammatica, S. *Phys. Rev. Lett.* **37**, 1360 (1976).
26. Pfister, G. *Phys. Rev. Lett.* **36**, 271 (1976).
27. Pfister, G. & Scher, H. *Adv. Phys.* **27**, 747 (1978).
28. Scher, H. & Montroll, E. W. *Phys. Rev.* **B12**, 2455 (1975).
29. Vardeny, Z., O'Connor, P., Ray, S. & Tauc, J. *Phys. Rev. Lett.* **84**, 1267 (1980).
30. Pai, D. M., & Enck, R. C. *Phys. Rev.* **B11**, 5163 (1975).
31. Williams, R. & Crandall, R. S. *RCA Rev.* **40**, 371 (1979).
32. Enck, R. C. & Pfister, G. *Photoconductivity and Related Phenomena* (eds Mort, J. & Pai, D.) Ch. 7 (Elsevier, Amsterdam, 1975).
33. Inokuchi, H. *Photoconductivity and Related Phenomena* (eds Mort, J. & Pai, D.) Ch. 5 (Elsevier, Amsterdam, 1975).
34. Onsager, L. *Phys. Rev.* **54**, 554 (1938).
35. Chen, I. & Mort, J. *Appl. Phys. Lett.* **37**, 952 (1981).
36. Meyerson, B. & Smith, F. W. *Solid St. Commun.* **34**, 531 (1980).
37. Pfister, G. & Morgan, M. *Phil. Mag.* **B41**, 209 (1980).
38. Abkowitz, M. & Enck, R. C. *J. Non-Cryst. Solids* **35-36**, 831 (1980).
39. Abkowitz, M. *The Physics of Selenium and Tellurium* (eds Gerlach, E. & Grosse, P.) (Springer, Berlin, 1979).
40. Troup, A., Mort, J., Grammatica, S. & Sandman, D. J. *J. Non-Cryst. Solids* **35-36**, 151 (1980).
41. Mott, N. F. *J. Non-Cryst. Solids* **35-36**, 1321 (1980).
42. Mort, J. et al. *Appl. Phys. Lett.* **38**, 277 (1981).

ARTICLES

Resonant structure of the asteroid belt

Stanley F. Dermott & Carl D. Murray

Center for Radiophysics and Space Research, Space Sciences Building, Cornell University, Ithaca, New York 14853, USA

Statistical techniques are applied to asteroid orbital data in an attempt to define the characteristics of the Kirkwood gaps in the distribution of the asteroids. A significant tendency is found for the eccentricities and inclinations of asteroid orbits to increase away from the gaps. These effects are observed throughout the distribution of main-belt asteroids. They are not merely confined to the immediate vicinities of the resonances and are not caused by observational selection effects. There is no evidence that the magnitude-frequency distribution of the asteroids changes near the gaps. Evidence is given for the existence of gaps in Hirayama families, possibly of a different nature from the classical Kirkwood gaps at locations corresponding to high order commensurabilities with Jupiter. These results support the theory that the Kirkwood gaps were formed by gravitational processes acting on individual asteroids throughout their lifetimes.

ALTHOUGH gaps in the distribution of the asteroid population are generally accepted to be associated with resonances with Jupiter¹, the exact mechanism which produces the gaps remains in dispute. The four major hypotheses² proposed to explain these Kirkwood gaps¹ are: (1) the statistical hypothesis (the proposal that asteroids are librating about the gaps and are therefore only rarely seen crossing the exact resonances), (2) the gravitational hypothesis (the gaps are formed by purely gravitational interactions with Jupiter), (3) the collisional hypothesis (the gravitational effects of Jupiter increase collision rates within the gaps, causing a wearing down or removal of material) and (4) the cosmogonic hypothesis (the gaps represent regions where asteroids failed to form during the early history of the Solar System). Although there are problems associated with each hypothesis, recent work³⁻⁶ has provided convincing evidence against the statistical hypothesis. It is important, therefore, to recognize which properties of the Kirkwood gaps could help to distinguish between the gravitational, collisional and cosmogonic hypotheses. Little work has been carried out in this field partly because of the failure to apply established statistical methods to the wealth of data available.

Resonance

The integral with respect to time of the expansion of the disturbing function for the gravitational interaction between two bodies (see, for example, ref. 7) contains divisors which become small when the ratio of the mean motions is close to a simple fraction of the form $p/(p+q)$, where p and q are integers. Such locations in the asteroid belt correspond to resonances with Jupiter and q is called the order of the resonance. The strength of a given resonance depends on several factors which include the distance between the bodies and the value of q : a large value of q implies a weak resonance. In the present study, we include all resonances with $q \leq 6$ which lie between the 1/3 and 1/2 resonances with Jupiter. These are the 1/3, 3/8, 2/5, 3/7, 4/9, 5/11 and 1/2 resonances. The limit $q \leq 6$ is arbitrary but reasonable. If Kirkwood gaps do have any characteristics other than a lack of asteroids, one would expect them to be more prominent in the vicinity of the strongest resonances.

Asteroids outside the 1/2 resonance are omitted. The number of asteroids between this resonance and Jupiter is small and some resonances in this region are associated with groupings of asteroids rather than gaps (see ref. 6). Asteroids inside the 1/3 resonance are also omitted because their orbits may have been influenced by Mars as well as by Jupiter⁸. Furthermore, the

resonances in this region, which is distant from Jupiter, are weak.

The *Tucson Revised Index of Asteroid Data* (TRIAD) is a convenient source of information on the orbital elements and physical properties of asteroids. In particular, Williams⁹ has calculated sets of proper elements (orbital elements after the effects of short-period oscillations have been removed) for 1,687 asteroids. Hirayama¹⁰ recognized certain groupings in the distribution of proper elements which he called families. The inclusion in our study of large numbers of asteroids whose orbital elements are similar or correlated for reasons other than those associated with resonance would give misleading results. We therefore exclude any asteroid which is designated a family member in the TRIAD file.

Thus, our study is confined to main-belt, non-family asteroids with semi-major axes a' in the range 2.50–3.28 AU whose proper elements are listed in the TRIAD file. 757 asteroids satisfy these requirements.

Statistical method

For each asteroid in our chosen population we calculate R , the ratio of the mean motions of Jupiter and the asteroid ($R < 1$). For each value of R we find the fractions R_2 and R_1 in the set 1/3, 3/8, 2/5, 3/7, 4/9, 5/11, 1/2, which bound R most closely from above and below. Then, using a method similar to that of Dermott¹¹, we define a parameter δ by

$$\begin{aligned} A &= (R - R_1)/(R_2 - R_1) \\ \delta &= 2(A - B) \\ B &= 0 \text{ if } A \leq 1/2 \\ B &= 1 \text{ if } A > 1/2 \end{aligned} \quad (1)$$

Thus, $-1 \leq \delta \leq +1$ and is a measure of the displacement of the asteroid from the nearest strong resonance. $\delta = 0$ corresponds to exact resonance and $|\delta| = 1$ is the greatest displacement from resonance that an asteroid can have.

Clearly the number density of asteroids systematically increases and decreases to a variable degree between the strong resonances. By defining δ we can superpose distributions to obtain a single distribution which is characteristic of the population as a whole. The relationship between δ and any other property of the asteroids or their orbits, such as their absolute magnitudes, types, diameters or proper eccentricities, can then be analysed.

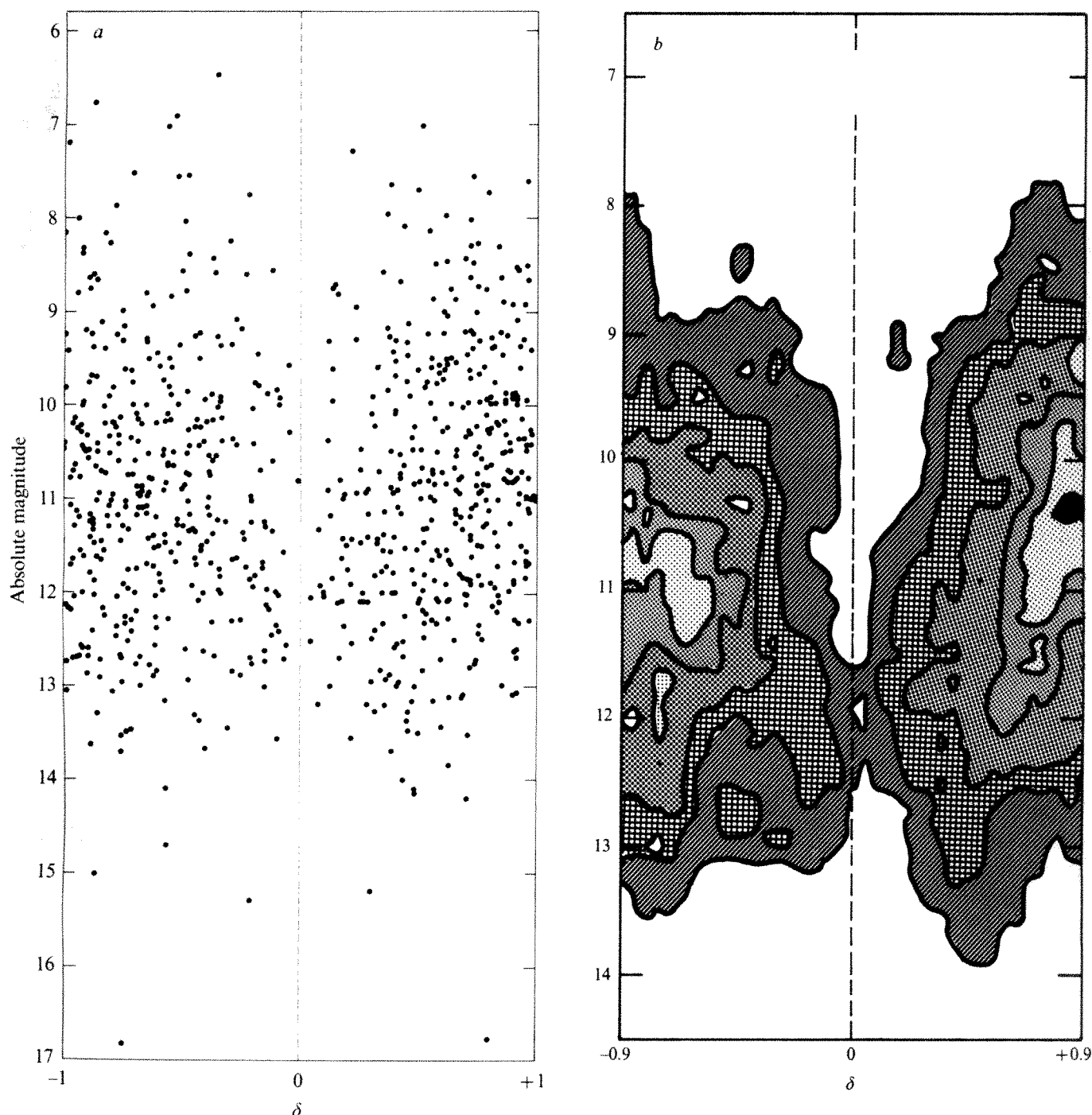


Fig. 1 *a*, A direct plot of absolute magnitude against δ (a measure of the displacement from the nearest strong resonance) for the 757 asteroids in the TRIAD file which have semi-major axes in the range 2.50–3.28 AU and which are not members of recognized families. The apparent lack of asteroids with large absolute magnitudes is an observational selection effect. *b*, A smoothed number density contour plot obtained by applying running box methods to the data shown in *a*. Note how the number density increases continuously as $|\delta|$ increases: there is no plateau. This implies that the effects of resonance pervade the whole distribution of main-belt asteroids. For absolute magnitude ≤ 10.3 the number density decreases as the absolute magnitude decreases—that is, there are relatively few asteroids of low magnitude. It is a combination of these two trends which results in the general V-shape of the contour plot.

Kirkwood gaps

Figure 1*a* plots absolute magnitude, $B(1, 0)$, against δ for all the asteroids in our chosen population. The absence of objects near $\delta = 0$, which corresponds to exact resonance, is the only accepted characteristic of Kirkwood gaps. It is useful to study a contour plot of the data shown in Fig. 1*a* and this is shown in Fig. 1*b*, which was obtained by applying a running box method to our data set. Note that as $\delta \rightarrow \pm 1$ the number density monotonically increases: the effects of resonance pervade the whole distribution of main-belt asteroids.

It has been suggested (see, for example, ref. 12) that larger objects avoid the gaps more than the smaller objects, and on first

examination Fig. 1*b* seems to confirm this. However, a combination of a decrease in number density with decreasing magnitude and an increase in number density with increasing $|\delta|$ would always give the illusion that low-magnitude objects avoid the gaps, even if the asteroids were in fact randomly distributed on either side of the resonances. Alternative methods are needed to settle this question.

By ordering our chosen population in absolute magnitude and calculating running means, we can plot the mean value of $|\delta|$ against the mean absolute magnitude (Fig. 2*a*). In Fig. 2*a*, and also Figs 2*b* and 3, the running box always contains 100(=N) asteroids and was shifted through the chosen population with a step length of 10 asteroids.

The central limit theorem (see, for example, ref. 13) relates σ , the standard deviation of the distribution of the means of the samples, to σ_p , the standard deviation of the chosen population by

$$\sigma = \frac{\sigma_p}{\sqrt{N}} \quad (2)$$

where N is the number of objects in each sample.

Figure 2a shows that $\langle |\delta| \rangle$ increases as absolute magnitude decreases. The result is significant at $\geq 3\sigma$ level. Thus, larger objects in our chosen population do avoid the gaps more than the smaller objects. However, by ordering our chosen population by sine of proper inclination, we can also show that proper inclination increases as $\langle |\delta| \rangle$ increases (Fig. 2b). Again, the result is significant at $\geq 3\sigma$ level. Furthermore, the data shown in Fig. 2 seem to be negatively correlated, and this is confirmed in Fig. 3, which plots the mean absolute magnitude against mean sine proper inclination.

That mean absolute magnitude should decrease as mean proper inclination increases is a consequence of observational selection¹⁴. Faint asteroids have a greater probability of discovery if their inclinations are low and close to the ecliptic. Thus, although the trends shown in Fig. 2 are undoubtedly real, it is probable that only one of them is a fundamental property of

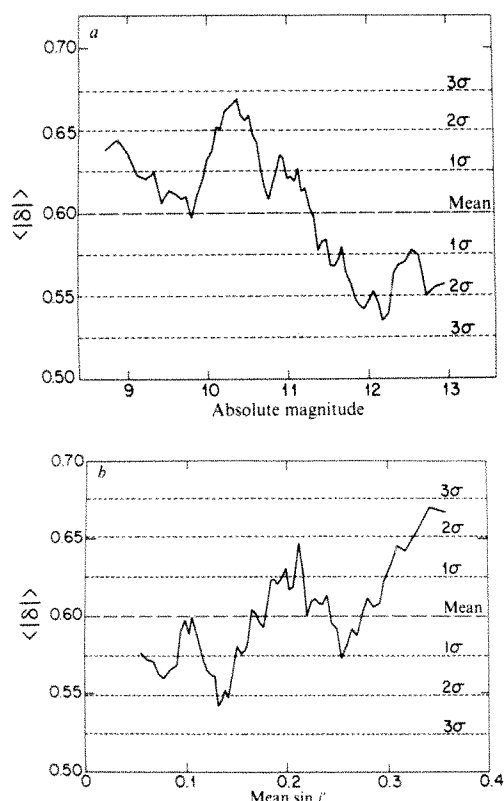


Fig. 2 A plot of the mean value of $|\delta|$ against: *a*, the mean absolute magnitude; and *b*, the mean sine proper inclination obtained by applying a running box method to the chosen population of 757 non-family asteroids which were sorted in order of increasing *a*, absolute magnitude; and *b* sine proper inclination. The σ levels are derived by the application of the central limit theorem to the distribution of the means of the samples of the population. The running box (and hence each sample) contained 100 asteroids and was stepped through the population 10 asteroids at a time. *a* Shows that low-magnitude objects do tend to have larger values of $|\delta|$, that is, that low-magnitude objects do avoid the gaps to a greater degree than do the high-magnitude objects. *b* Shows that there is a tendency for displacement from resonance to increase with inclination. In both *a* and *b* the result is significant at $\geq 3\sigma$ level.

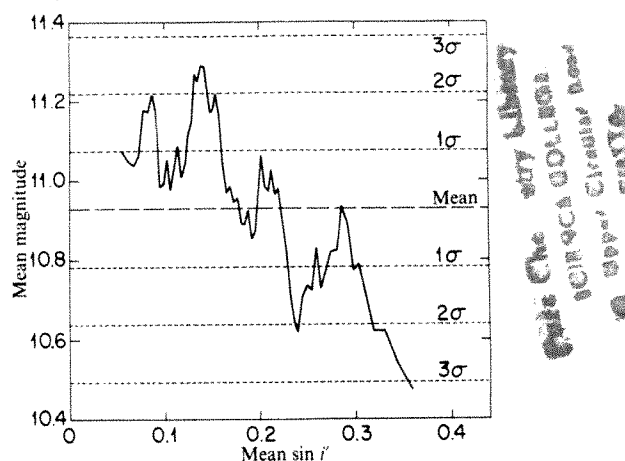


Fig. 3 A plot of the mean value of absolute magnitude against mean sine proper inclination obtained by applying a running box method to our chosen population of 757 non-family asteroids which were sorted in order of increasing sine proper inclination. The correlation is consistent with the effects of observational selection which produce an excess of high magnitude, low inclination objects¹⁴.

Kirkwood gaps and the result of resonance rather than observing procedures. We must now determine which trend is the fundamental one, and we do this by analysing a data set which is free of all observational selection effects.

Bias-free statistics

On the basis of results from the McDonald survey¹⁵, Kiang¹⁴ estimated that the catalogued asteroid population is 95% complete up to mean opposition magnitude, $B(a, 0) = 15$. Mean opposition magnitude is related to absolute magnitude by

$$B(a, 0) = B(1, 0) + 5 \log a(a - 1) \quad (3)$$

where a is the semi-major axis of the asteroid in AU. Furthermore, Kiang points out that of the 587 new objects detected in the survey only two had apparent magnitudes < 14 . To ensure complete elimination of all observational selection effects, even at the risk of being over-conservative, we adopt as an upper limit that value of absolute magnitude which corresponds to mean opposition magnitude 14 at the largest semi-major axis of our chosen population, 3.25 AU. This gives us (see equation (3)) the conditions

$$\begin{aligned} B(1, 0) &< 9.68 \\ B(a, 0) &< 14.00. \end{aligned} \quad (4)$$

Of our chosen population of 757 asteroids, 144 satisfy the above conditions. These asteroids constitute our bias-free set.

Using the bias-free set we have repeated the plot of $\langle |\delta| \rangle$ against $\langle B(1, 0) \rangle$ and have found no significant trend. The set yields no evidence that the magnitude-frequency distribution of the asteroids changes near the Kirkwood gaps. However, by repeating the plot of $\langle |\delta| \rangle$ against $\langle \sin i' \rangle$ we have found that the trend shown in Fig. 2b persists. Thus, we conclude that it is the latter which is of fundamental importance and that the trend shown in Fig. 2a is merely a consequence of observational selection effects.

We have also discovered that, in addition to the proper inclinations i' , the proper eccentricities e' also tend to increase away from the gaps. Here, these trends are best illustrated by Fig. 4 in which we plot $\langle V \rangle$ against $\langle |\delta| \rangle$, where

$$V = (e'^2 + \sin^2 i')^{1/2} \quad (5)$$

We consider it significant that the trend shown in Fig. 4 does not level off for large values of $|\delta|$, suggesting that the effects of resonance are not confined to the immediate vicinities of the resonances, but pervade the whole distribution of main-belt asteroids.

Themis family

Franklin *et al.*¹⁶ have recently suggested that the Kirkwood gaps were formed early in the history of the Solar System at a time when the asteroids moved in near-circular, near-coplanar orbits and low-velocity collisions were frequent. They consider that the present high values of inclinations and eccentricities of orbits in the asteroid belt were produced later by some process unrelated to gap formation. We can now see that their hypothesis does not account for the characteristic trends in i' and e' which we have discovered. Further evidence contrary to their hypothesis may be provided by the Hirayama families.

The 144 asteroids that constitute our bias-free set are among the largest asteroids in the main belt. Because the lifetime of an asteroid against collisions is proportional to $\sqrt{D}$, where D is the asteroid diameter¹⁷, this set also consists of some of the oldest asteroids. It would be of interest, therefore, to look for gaps in the distributions of objects which are probably relatively recent additions to the asteroid population. The larger asteroid families are generally believed^{18,19} to be the results of collisional disruption of single bodies. The Eos family is estimated²⁰ to have been formed $\geq 10^6$ yr ago, while the Themis family is thought to be somewhat older because of the larger spread of its orbital elements. A large spread in semi-major axes increases the likelihood of sufficiently strong resonances occurring within a given family. The Themis family, with 62 catalogued members⁶ lying between 3.08 and 3.18 AU, provides the best set of relatively young asteroids in which to search for gaps associated with resonances.

There are five resonances with $q \leq 11$ located within this family. When we apply a running box method to the distribution of a , we find gaps associated with the two strongest resonances (Fig. 5). In addition, the family seems to be bounded on one side by the 5/11 resonance (3.075 AU). One may question whether these gaps are real, in the sense that they represent a genuine absence of material or whether, because family members are identified by clustering of orbital elements, they represent the failure to identify asteroids in these regions as family members because their eccentricities and inclinations, but not, perhaps, their semi-major axes, have been altered by the effects of resonance. The answer to this important question may be obtained from more complete studies of background asteroids in

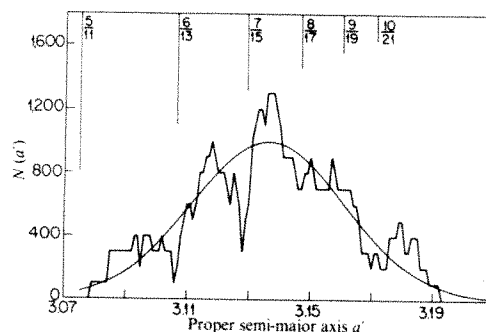


Fig. 5 The distribution of the 62 asteroids in the Themis family as a function of proper semi-major axis. The plot was obtained using a running box method with a box length of 0.01 AU and a shift of 0.001 AU between data points. $N(a')$ is the number of asteroids per unit AU contained in each box. The smooth curve represents the best gaussian fit to the data. A gaussian distribution is characteristic of catastrophic disruption²¹. The location and approximate relative strength of all resonances up to and including order 11 which lie in the region of the Themis family are shown. The family appears bounded to the left by the 5/11 resonance. The strong 1/2 resonance lies at 3.28 AU. The two strongest resonances which lie in the Themis region (6/13 and 7/15) coincide with gaps in the distribution. These gaps may be of a similar nature to the Kirkwood gaps or they could be regions where the resonances have altered the proper elements of asteroids causing them to be overlooked as potential Themis family members. Observations of the physical properties of background asteroids in these regions should help to distinguish between these hypotheses.

these gaps to determine whether or not they have the physical characteristics of Themis members.

On the basis of our studies of the Themis and other families, in particular, the Eos and Koronis families, we tentatively conclude that the gaps in their distributions of semi-major axes are real and thus that some Kirkwood gaps have been formed since the time of formation of the Solar System.

Conclusions

When the bias in the data due to observational selection is removed, there is no significant tendency for low-magnitude objects to be further away from the resonances than the high-magnitude objects. Such a magnitude effect, if found, would have supported the collisional hypothesis for the origin of the gaps.

There is a significant tendency for objects further away from the resonances to have larger eccentricities and inclinations. The process responsible for the formation of the Kirkwood gaps seems preferentially to have removed those objects near the resonances which had large values of e' and i' . Since the magnitude of the gravitational disturbing function acting on an asteroid increases with the eccentricity and the inclination of the asteroid, our results strongly suggest that it is the gravitational effect of Jupiter, acting on individual asteroids, which is responsible for the formation of the Kirkwood gaps. It is possible that such a mechanism has produced the gaps over the lifetime of the Solar System, but further studies of Hirayama families should provide more accurate measurements of the timescale of such a process.

The Kirkwood gaps are not simply regions of small asteroid number density, neither are they restricted to the immediate vicinities of the resonances, for the effects of resonance seem to pervade the entire belt of asteroids between the 1/3 resonance and Jupiter itself. The characteristics we have described should place severe constraints on any theory of the formation of the Kirkwood gaps. They also challenge our understanding of the dynamical stability of the Solar System.

We thank P. Goldreich, J. Gradie and T. C. Van Flandern for helpful discussions and B. Zellner for providing us with a copy of the TRIAD file. This research was supported by NSF grant AST 79164-74-A01.

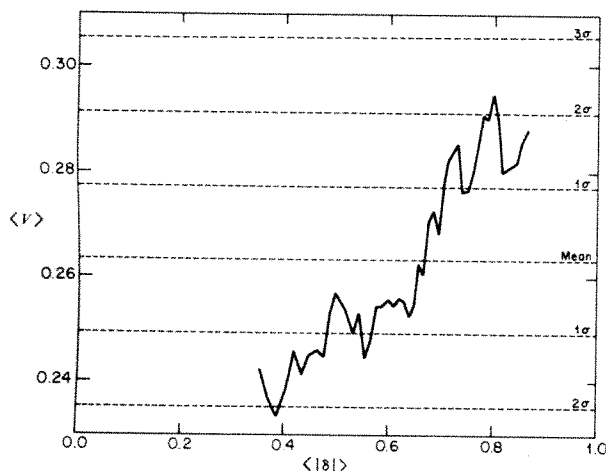


Fig. 4 A plot of the mean value of $V = (e'^2 + \sin^2 i')^{1/2}$ against the mean value of $|\delta|$ obtained by applying a running box method to the bias-free set of 144 non-family asteroids which were sorted in order of increasing $|\delta|$. For this plot, the running box contains 50 asteroids and there is a shift of two asteroids between data points. Note that the plot has no plateau: the trend is present throughout the distribution. This tendency for both e' and i' to increase with increasing $|\delta|$ is a fundamental property of the Kirkwood gaps and is indicative of the effects of gravitational forces acting on individual asteroids as a possible mechanism for the origin of the Kirkwood gaps.

Received 22 December 1980; accepted 5 February 1981.

1. Kirkwood, D. *Meteoritic Astronomy* (Lippincott, Philadelphia, 1867).
2. Greenberg, R. & Scholl, H. in *Asteroids* (ed. Gehrels, T.) 310–333 (University of Arizona Press, 1979).
3. Message, P. J. in *The Theory of Orbits in the Solar System and in Stellar Systems* (ed. Contopoulos, G.) 197–222 (Academic, London, 1966).
4. Schweizer, F. *Astr. J.* **74**, 779–788 (1969).
5. Wiesel, W. E. *Celestial Mech.* **13**, 3–37 (1976).
6. Sinclair, A. T. *Mon. Not. R. astr. Soc.* **142**, 289–292 (1969).
7. Brouwer, D. & Clemence, G. M. *Methods of Celestial Mechanics* (Academic, New York, 1961).
8. Williams, J. G. thesis, Univ. California (1969).
9. Williams, J. G. in *Asteroids* (ed. Gehrels, T.) 1040–1063 (University of Arizona Press, 1979).
10. Hirayama, K. *Proc. phys.-math. Soc. Japan Ser. II*, **9**, 354–361 (1918).
11. Dermott, S. F. *Nature phys. Sci.* **244**, 18–21 (1973).
12. Zellner, B. & Bowell, E. in *Comets, Asteroids, Meteorites: Interrelations, Evolution and Origins* (ed. Delsemme, A. H.) 185–197 (University of Toledo, 1977).
13. Lindgren, B. W. *Statistical Theory* (Macmillan, New York, 1960).
14. Kiang, T. *Icarus* **5**, 437–449 (1966).
15. Kuiper, G. P. et al. *Astrophys. J. Suppl.* **3**, 289–428 (1958).
16. Franklin, F. A., Lecar, M., Lin, D. N. C. & Papaloizou, J. *Icarus* **42**, 271–280 (1980).
17. Chapman, C. R. in *Asteroids* (ed. Gehrels, T.) 25–60 (University of Arizona Press, 1979).
18. Gradie, J. & Zellner, B. *Science* **197**, 254–255 (1977).
19. Gradie, J. thesis, Univ. Arizona (1978).
20. Steins, K. A. *Astr. Zh.* **33**, 246–250 (1956).
21. Wiesel, W. *Icarus* **34**, 99–116 (1978).

Interacting genes in nematode dauer larva formation

Donald L. Riddle, Margaret M. Swanson & Patrice S. Albert

Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211, USA

The dauer larva of Caenorhabditis elegans is a developmentally arrested stage induced by starvation or overcrowding. Mutant genes controlling the ability to form dauer larvae interact in a way which allows them to be ordered in a pathway. Mutant phenotypes suggest that the pathway corresponds to neural processing of environmental stimuli.

To the extent that the development of organisms is controlled by genes which act in sequence¹, it should be possible to determine a hierarchy of genes affecting the expression of certain traits. What follows is an account of a study, in the language of classical genetics, of the hierarchical relationship between the genes involved in the formation of the so-called dauer larval stage of the free-living soil nematode *Caenorhabditis elegans*. The genetic and anatomical simplicity of the organism^{2–5} fit it for such a study. Our work shows, however, that the expected hierarchy is surprisingly intricate.

Larvae of *C. elegans* undergo four moults (Fig. 1), at which the cuticle is replaced, but at the second may enter the 'dauer' or 'persistent' (German: *dauer* = hard, enduring) state⁶ in response to starvation or overcrowding. Thus arrested in development but specialized for survival and dispersal⁷, dauer larvae survive exposure to detergents and other chemicals⁸, do not feed because their mouths are plugged (Fig. 2) and may survive for months until they encounter food, moult and resume development. The duration of the dauer stage seems not to affect post-dauer longevity, suggesting that ageing is interrupted⁹.

Table 1 Suppression of dauer-constitutive mutations by dauer-defectives

Dauer defective mutants	Multiply-marked dauer-constitutive strains						
	DR194 (<i>daf-11 V</i>)	DR189 (<i>daf-8 I</i>)	DR133 (<i>daf-7 III</i>)	DR202 (<i>daf-14 IV</i>)	DR204 (<i>daf-1 IV</i>)	CB2196 (<i>daf-4 III</i>)	DR130 (<i>daf-2 III</i>)
CB1376 (<i>daf-3 X</i>)	+	+	+	+	+	+	–
CB1386 (<i>daf-5 II</i>)	+	+	+	+	+	+	–
DR20 (<i>daf-12 X</i>)	+	+	+	+	+	+	–
DR25 (<i>daf-20 X</i>)	+	+	+	+	+	–	+
DR26 (<i>daf-16 I</i>)	+	+(60%)	+(20%)	+(25%)	+(35%)	–	+
CB1377 (<i>daf-6 X</i>)	+(70%)	+(30%)	+	+(35%)	–	–	–
DR27 (<i>daf-17 I</i>)	+	+	+	–	–	–	+
CB1375 (<i>daf-18 IV</i>)	+	+	–	–	–	–	–

+ Indicates that, in a particular constitutive-defective double mutant, the constitutive phenotype is suppressed (the defective mutation is epistatic). – Indicates gene pairs in which the constitutive is epistatic (no suppression of the constitutive phenotype). In cases where suppression is not complete (*daf-6* and *daf-16*), the fraction of the double-mutant population which escapes dauer formation is given in parentheses as per cent suppression. As the only available mutant alleles of these genes are slightly leaky dauer-defectives, a low residual level of gene activity may account for reduced efficiency of suppression. Dauer-defective strains are all single mutants with allele numbers corresponding to the strain numbers given. Gene names and linkage groups are given in parentheses. The multiply-marked dauer-constitutive strains have the following genotypes, where allele numbers are given in parentheses after the corresponding gene name: DR194, *dpy-11(e224)daf-11(m47)unc-39(e257)V*; DR189, *dpy-5(e61)daf-8(e1393)unc-29(e1072)II*; DR133, *dpy-1(e1)daf-7(e1372)unc-32(e189)III*; DR202, *unc-24(e138)daf-14(m77)IV*; DR204, *daf-1(m40)dpy-4(e1166)IV*; CB2196, *daf-4(e1364)unc-32(e189)III*; DR130, *dpy-1(e1)daf-2(e1370)unc-32(e189)III*. Because it was impossible to predict in advance if a double mutant would be constitutive or defective (or even viable), closely linked visible markers were used to identify the presence or absence of *daf* alleles in newly constructed strains. To construct the strains, males homozygous for a dauer-defective (suppressor) mutation were crossed with *is* dauer-constitutive hermaphrodites carrying closely linked, flanking mutations. Heterozygous F₁ hermaphrodites produced in these crosses were wild type in phenotype, and segregated dumpy, uncoordinated animals by 'selfing' at 15 °C. The dauer-constitutive phenotype is not expressed at 15 °C. In each test, 15 marked segregants were incubated at 25 °C to determine if suppression occurred. If the suppressor gene is unlinked, three-quarters of the marked animals are expected to carry the suppressor gene in either the heterozygous or homozygous state. All such animals segregate homozygous progeny which, in the case of suppression, will grow to maturity at 25 °C. If there is no suppression, all progeny become dauer larvae. In cases where closely linked genes were tested, recombinants carrying both mutations were selected by using an intervening visible marker as in the case of *daf-8* and *daf-16*. Starting with a heterozygote (+ + *daf-16/daf-8 unc-29* +), a *daf-8 daf-16* recombinant was selected by picking F₁ progeny which are dauer-constitutive, but not uncoordinated (*daf-8 + daf-16/daf-8 unc-29* +). At 15 °C, this heterozygote segregated the double homozygote required to test the epistatic relationship of *daf-8* and *daf-16* at 25 °C. Crosses with the *daf-10* mutant are not included here. Appropriate double mutants were constructed by another method¹², because homozygous *daf-10* males do not mate¹¹. All the constitutive mutations except *daf-4* have been combined with *daf-10*, and each failed to be suppressed.

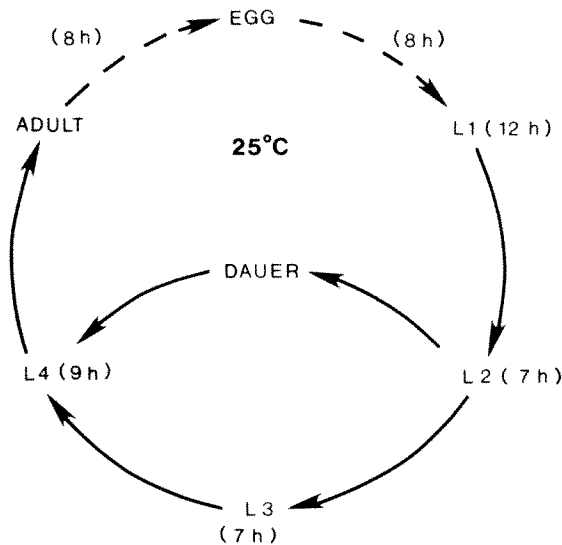


Fig. 1 The life cycle of *C. elegans*. The duration of each developmental stage during growth at 25°C on Petri plates seeded with *Escherichia coli* strain OP50 (ref. 2) is given in parentheses. The solid arrows represent moults. Moulting times were determined by monitoring pharyngeal pumping in synchronized populations¹². Wild-type dauer larvae obtained from starved populations initiate pharyngeal pumping 2–3 h after being placed in food. The moult to the L4 stage occurs after an additional 8 h. The 8-h time from egg to L1 is the interval between egg laying and hatching. The entire life cycle requires 3 days at 25°C.

Such stages of arrested development occur in other nematodes, and the similarity with the obligate (usually infective) form of many pathogenic parasites should not be entirely overlooked¹⁰.

In our genetic studies, we have isolated mutants of *C. elegans* of two kinds—'defective' mutants which do not produce dauer larvae even when starved¹¹, and 'constitutive' mutants which produce dauer larvae even with abundant food¹². The genes concerned are named *daf* (for 'dauer larva formation').

We have screened clones for defective mutants by treating nematodes with the mutagen ethyl methane sulphonate (EMS) and selecting clones of F₂ hermaphrodite progeny which do not form dauer larvae after starvation; a total of at least 15 genes is involved^{11,13}. We have obtained 38 constitutive mutants from nematodes treated with either EMS or DES (diethyl sulphate) by selection for dauer larvae with 1% SDS¹⁴.

Thirty-three of the constitutive mutants are *ts* (temperature-sensitive), while five are 'leaky' (incompletely penetrant) at all growth temperatures. The best of the *ts* mutants grow normally at 15°C, but all larvae grown at 25°C become dauers at the second moult. The dauer larvae cannot recover until the incubation temperature is shifted down to 15°C. The *ts* mutants are all recessive, and they define seven complementation groups (see Table 1). SDS selection can only yield conditional or leaky mutants which are able to recover and reproduce. The isolation of unconditional (non-recovering) dauer-constitutive mutants will be described elsewhere.

Epistatic ordering of genes

A relationship between constitutive and defective mutations was detected by selecting revertants of *ts* dauer-constitutives, *daf-2 daf-4* and *daf-7* (Table 2). Homozygous revertants, arising in the F₂ generation after mutagenesis, grow and reproduce at 25°C, while parental-type larvae become dauers and stop growing. Most revertants not only escape dauer formation, but they fail to form dauers even when starved. Crossing the revertants with the parental strain revealed that they were double mutants which carried not only the parental constitutive mutation, but also were homozygous for an epistatic, recessive dauer-defective mutation. Thus, some dauer-defective mutations can act as suppressors of dauer-constitutives.

Twenty-five independently isolated revertants were distributed among six complementation groups (Table 2) and some

revertants were alleles of dauer-defective genes which had already been genetically mapped (Fig. 3). To determine which dauer-constitutive mutants could be suppressed by various defectives, we constructed a series of multiple mutants by genetic crosses (Table 1). One interpretation of these data could be that constitutive and defective mutations produce competitive influences on dauer larva formation in a double mutant. The more extreme alleles might be epistatic to less extreme alleles, so that the cross-suppression tests rank the mutations quantitatively in order of severity. Several observations argue against this hypothesis as a general explanation for the observed interactions. First, there is no uniform gradient of suppression. For example, the *daf-12* mutation suppresses *daf-4* but not *daf-2* while the *daf-20* mutation exhibits the opposite pattern. Second, ranking the *ts* mutants by the degree of expression at intermediate temperatures does not produce a correspondence to the suppression pattern¹². Based on the degree of expression of the *ts* constitutive strains at a range of temperatures between 15 and 25°C, the *daf-7* strain was judged to carry the most severe *ts* allele, while the *daf-14* strain carries the least severe allele. Third, *daf-16* provides an absolute block to the most

Table 2 Origin of dauer-defective mutations

Gene and linkage group	Reference allele	No. of alleles collected*			Estimated mutation frequency
		By reversion†	Other‡		
<i>daf-10 IV</i>	<i>e1387</i>	—	—	2	5×10^{-4}
<i>daf-18 IV</i>	<i>e1375</i>	—	—	1	
<i>daf-17 I</i>	<i>m27</i>	—	—	5	1×10^{-4}
<i>daf-6 X</i>	<i>e1377</i>	—	—	1	
<i>daf-16 I</i>	<i>m26</i>	—	—	1	2×10^{-5}
<i>daf-20 X</i>	<i>m25</i>	—	—	2	5×10^{-5}
<i>daf-3 X</i>	<i>e1376</i>	5	2	—	3×10^{-4}
<i>daf-5 II</i>	<i>e1386</i>	7	2	—	5×10^{-4}
<i>daf-12 X</i>	<i>m20</i>	1	—	—	5×10^{-5}

* Complementation tests were performed by crossing homozygous or hemizygous dauer-defective males with dumpy (*dpy*) dauer-defective (*daf*) hermaphrodites. Starved populations derived from cloned, non-dumpy F₁ progeny were scored for the presence or absence of dauer larvae both visually and by SDS treatment^{11,13}. Populations arising from heteroallelic F₁ progeny do not form dauers. Complementation *daf* mutations, on the other hand, allow many dauers to be formed even though some homozygous dauer-defective progeny are segregated in the F₂ generation. Because homozygous *daf-10* males do not mate¹¹, heterozygous *daf-10/+* males were crossed with *dpy daf* hermaphrodites. In these crosses, half the F₁ progeny produce dauer-defective populations when the *daf* mutations are allelic.

† To obtain revertants of *ts* dauer-constitutive mutations, strains CB1372, *daf-7(e1372)III*; CB1364, *daf-4(e1364)III*; CB1370, *daf-2(e1370)III*; CB1368, *daf-2(e1368)III* and CB1391, *daf-2(e1391)III* were mutagenized with EMS (ref. 6). In each experiment 4 adult hermaphrodites were placed on each of 50 100-mm diameter Petri plates seeded with OP50 and allowed to produce about 200 F₁ progeny per plate. After an initial 6 days of incubation at 15°C, adult F₁ progeny were shifted to 25°C for an additional 4 days. Adult F₂ revertants were cloned and re-tested for growth at 25°C, and for the dauer-defective phenotype after starvation. To assure independence of the mutations, no more than one revertant clone from each plate was saved. Homozygous revertant males were backcrossed to *dpy daf* hermaphrodites. The appearance of dauer-constitutive cross progeny in the F₁ demonstrates the continued presence of the parental dauer-constitutive allele in the revertant strain, as well as the recessive character of the suppressing dauer-defective mutation. The dauer-constitutive alleles were then removed from the genetic background of each revertant by crossing the homozygous revertant males with strains carrying *dpy* markers closely linked to the constitutive locus. Doubly homozygous *dpy;daf* F₂ segregants which no longer carried the constitutive mutation were saved for complementation tests. The *dpy;daf* strains were again crossed with wild-type males to remove the *dpy* marker. In each case homozygous or hemizygous males derived from a non-dumpy, dauer-defective segregant were used for cross-suppression tests (see Table 1).

‡ Alleles *e1387*, *e1376* and *e1386* were obtained by cloning F₂ progeny of mutagenized N2 hermaphrodites as described elsewhere¹¹. All mutants were subsequently backcrossed and re-segregated before complementation and genetic mapping. The *daf-18(e1375)III* mutation was discovered in strain CB1062, *vab-3 X*, and removed from the *vab-3* background by crossing with wild-type males and re-segregating the unlinked *e1375* allele. The *e1377* mutant was isolated after EMS mutagenesis by Dr James A. Lewis. The frequency of mutation in each gene in standard conditions of EMS mutagenesis² was estimated by dividing the total number of mutants obtained (columns 3–6) by twice the number of F₁ broods examined.

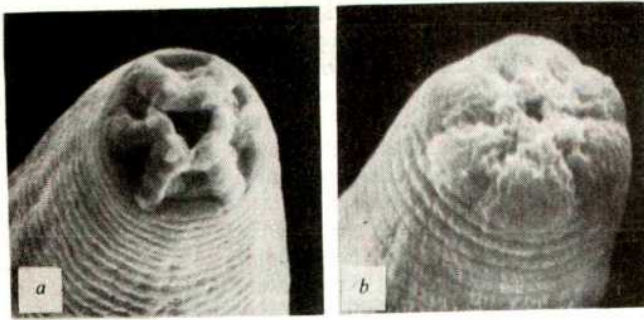


Fig. 2 External morphological differences between an L2 larva (a) and a dauer larva (b) are shown by scanning electron micrographs of the nose. The field width in each micrograph is $7.2\ \mu\text{m}$. An internal plug which completely closes the mouth of the dauer (b) is revealed in transmission electron micrographs of transverse sections (not shown). The dauer larva is thinner than the L2, and has an increased specific gravity due to radial shrinkage of the body at the dauer-specific moult⁷. In preparation for scanning electron microscopy, larvae were fixed overnight in cold 3% glutaraldehyde in 0.1 M phosphate buffer, pH 7.1, and post-fixed for 6 h in cold 1% buffered OsO_4 . Specimens were dehydrated in a graded series of ethanol, critical point dried in CO_2 , mounted on copper adhesive tape and gold coated. Photographs were taken at 10 kV on a Jeol JSM-1 scanning electron microscope.

extreme (non-recovering) constitutive mutations, although it does not suppress *ts* alleles of certain constitutive genes (in preparation).

A functional sequence

The data suggest that there is an order to the functions controlled by the *daf* genes, and that the constitutive and defective mutations affect a common sequence. On the basis of the suppression patterns (Table 1) the pathway in Fig. 4 can be constructed with individual genes assigned to the control of sequential steps. If the constitutive mutants were simply blocked in entry to the L3 stage (see Fig. 1), while defective mutants were blocked in a separate sequence to the dauer larva, the constitutive-defective double mutants could not continue larval development at all. Instead, the data indicate that the dauer-defective mutants are blocked in the pathway to the dauer stage while the dauer-constitutives generate a false, internal signal which causes the mutants to form dauer larvae even in the absence of appropriate environmental cues. If the pathway is blocked after the false signal, a double mutant will be dauer-defective. However, if the false signal is generated after the block, a double mutant will be dauer-constitutive.

An unambiguous order of function can be assigned to most dauer-constitutive genes. However, *daf-2* seems to be in a distinct branch of the pathway. The *daf-16* and *daf-20* mutations suppress *daf-2* but not *daf-4*, while *daf-3*, *daf-5* and *daf-12* suppress all constitutives except *daf-2*. A block in the upper branch of the pathway is both necessary and sufficient to produce a dauer-defective phenotype. Therefore, we conclude that the lower pathway specified by *daf-2* does not function in these conditions. This second pathway may represent the response to different environmental signals, or it may function only in *daf-2* mutants.

Our goal is to establish specific behavioural or morphological correlates with the mutant steps in the pathway. Two steps, represented by *daf-10* and *daf-6*, are correlated with chemotaxis defects, and both mutants exhibit ultrastructural alterations¹¹ in specific neurones which have previously been implicated in the chemosensory response to salts^{15,16}. These dauer-defective mutants exhibit pleiotropic defects, both in chemosensory behaviour and in neuroanatomy. Nevertheless, ultrastructural evidence suggests that one or more amphidial neurones mediate a sensory signal for dauer larva formation¹¹. The amphids are a pair of sense organs consisting of 12 neurones, 8 of which extend afferent processes anteriorly to openings on the lateral sides of the mouth^{17,18}.

The genetic pathway may correspond to a neural pathway involved in detection of an environmental signal by sensory

receptors followed by transduction of the signal to neurosecretory cells. According to this hypothesis, neurosecretion would then alter the hormonal balance in the nematode to initiate dauer larva morphogenesis. The final step in the pathway may represent defects in the reception of the neurosecretory signal. Underlying the neural pathway would be the developmental processes which produce a functional set of nerve cells. For example, a mutation affecting morphogenesis of sensory neurones would affect the pathway at a position corresponding to that sensory function. Thus, we expect that many of the genes exert their primary effects on development well before the pathway is actually set in motion physiologically. This contention is supported by the results of temperature-shift and temperature-pulse experiments which define temperature-sensitive periods before the L2 stage for six of the *ts* dauer-constitutive mutants¹².

One or more steps in the pathway may involve amphid function. The *daf-6* mutant exhibits amphid defects as does *daf-10*, but other neurones in the *daf-6* mutant are also affected¹¹. None of the other mutations positioned in the pathway exhibits the behavioural abnormalities characteristic of the non-chemotactic dauer-defective mutants, *daf-6* and *daf-10*.

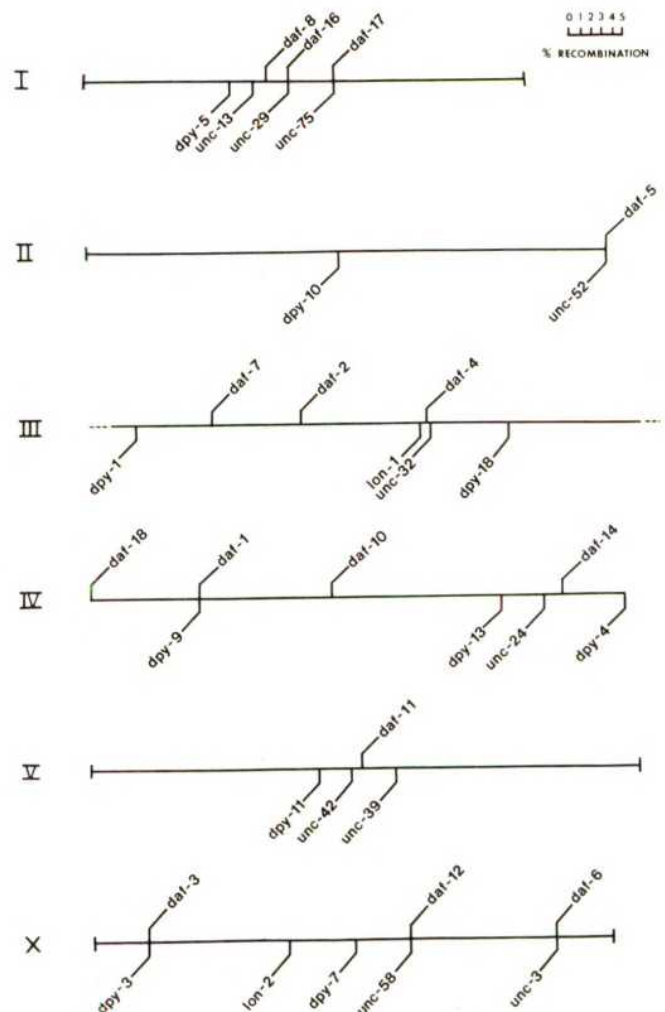
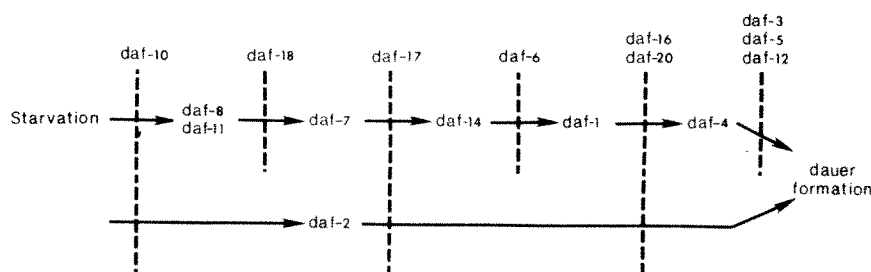


Fig. 3 Simplified genetic map of *C. elegans* showing the six linkage groups with *daf* loci labelled above the lines. Other markers used in mapping crosses and in cross-suppression tests (Table 1) are shown below the lines. Linkage relationships are as described previously²¹. Gene designations are *daf* (dauer larva formation), *dpy* (dumpy), *lon* (long) and *unc* (uncoordinated). The locations are shown only for mutations which have been unambiguously ordered. When two genes are joined at the same location, no recombination has been detected between them. The *daf-20* gene is X-linked, but is not shown because its location has not been determined.

Fig. 4 A simple genetic pathway for dauer larva formation, one of several possible representations, based on epistatic relationships between dauer-constitutive mutations and dauer-defective mutations. Gene names of dauer-constitutives are given in the pathway itself to represent points where false signals may be initiated. Mutations in dauer-defective genes block the pathway at the positions shown by dashed lines. The *daf-10* mutation is drawn to block both branches of the pathway on the basis of indirect evidence. First, both *daf-8* and *daf-2* are epistatic to *daf-10*, and second, the *daf-10* mutation blocks recovery of both *daf-2* and *daf-8* dauer larvae (see text).



However, if the gene order corresponds to a pathway of neurotransmission, electron microscopic analysis of *daf* mutants positioned near the end of the pathway may reveal anatomical defects in neurosecretory cells.

Genetic complexity of the pathway

C. elegans possesses three to four thousand essential genes^{2,19}. We estimate, based on mutation frequency and the fraction of new mutations which recur in previously characterized genes (in preparation), that there are ~100 genes affecting dauer larva formation. Dauer-constitutive mutants are rare, occurring at an aggregate frequency of 0.2% or about four times the mutation frequency per average gene in standard conditions of EMS mutagenesis^{2,13}. All seven of the constitutive genes detected by our mutant selection are represented by more than one allele. Therefore, we conclude that we are approaching saturation of the *ts* dauer-constitutive mutant class.

Two other experiments support this view. First, we have used dauer-defective mutants late in the pathway (*daf-16* and *daf-20*) as parent strains for selecting a restricted class of dauer-constitutive mutants which might define even later steps in the pathway (see Fig. 4). After mutagenesis with EMS or DES, the progeny of nearly 9,000 F₁ animals were screened. This population is of sufficient size to detect four or five alleles of the average gene in *C. elegans*. In fact, five new alleles of *daf-4* were found, but no new dauer-constitutive genes were identified. Second, mutants with maternal effects have not been overlooked. All eight alleles of *daf-1* share this property. Homozygous *daf-1* animals do not form dauer larvae if their maternal parent was heterozygous (+/*daf-1*). Selection for maternal-effect mutants (not expressed until the F₃ after mutagenesis) produced six new alleles of *daf-1*, but no other constitutive genes with a maternal effect.

Although the genetic evidence suggests that there may not be many more than 7 genes of the *ts* constitutive class, we estimate that there may be as many as 100 genes of the defective type¹³. It is likely that more defective genes will be positioned within the pathway, since four of the nine genes in the pathway are represented by only a single isolate (see Table 2) and other dauer-defective mutants have not been tested for their suppression properties. The frequency of dauer-defective mutants obtained by EMS mutagenesis of wild-type *C. elegans* is 5%. Many of these mutations may exert their effects indirectly as in the case of *daf-10* and *daf-6*, two mutations which affect neural development.

Other genes do not affect the signal, but do affect the subsequent process of dauer larva morphogenesis. For example,

a *daf-13* mutant forms apparently normal dauer larvae in response to starvation, but the dauers are sensitive to SDS treatment, possibly because of a defect in cuticle maturation (unpublished results). When the *daf-13* mutation is combined with the dauer-constitutive, *daf-2*, the constitutive phenotype is unaltered except that the constitutively formed dauer larvae are sensitive to SDS. We believe *daf-13* to be representative of a class of 'process' mutations which do not affect the signal pathway and do not exhibit epistatic interactions with dauer-constitutive mutations.

There are few, if any, genes which are involved specifically in exit from the dauer stage. Selection (by resistance to SDS) for mutants which are abnormally slow in dauer larva recovery produced no new mutant types. The only mutants found were dauer-constitutives (which recover slowly) and leaky dauer-defectives (affected in both entry and exit). This suggests that the genetic requirements for recovery from the dauer stage are similar to the requirements for entry. To determine if all dauer-defective genes specifying the entry pathway also affect exit from the dauer stage, we constructed the following three strains in which the constitutive gene, *daf-1*, is epistatic: *daf-1 daf-10*, *daf-1 daf-17* and *daf-1 daf-6*. In both the *daf-10* and *daf-6* double mutants, recovery of constitutively formed dauers (on temperature-shift to 15°C) is almost completely inhibited, showing that function of these genes is required for both entry and exit, but in the case of the *daf-1 daf-17* double mutant, dauer larva recovery is at most only marginally inhibited. Thus, the genes required for exit from the dauer stage may be simply a subset of the genes required for entry.

Dauer larva formation seems to be relatively unaffected by most of the animal's genes, allowing us to study this sequence in relative isolation from the rest of development. The properties of this system permit an unusual application of classical genetic methods to the analysis of a developmental pathway. Elucidation of the physiology of dauer larva formation will require further work to identify the specific molecules which act as environmental cues, the neurones involved in signal processing, what neurotransmitters these cells use and what hormones mediate the response. As a formal genetic problem, dauer larva development seems about as complex as that of bacteriophage T4 (ref. 20), and it is therefore a reasonable goal to characterize its genetic specification completely.

This work was supported by USPH grant HD11239 and National Research Service Award HDO5654. We thank Susan Brown for the scanning electron micrographs, Sydney Brenner, Robert Waterston and Gunter Stent for helpful discussion, and Betty Hearn for manuscript preparation.

Received 8 October 1980; accepted 13 January 1981.

1. Beadle, G. W. & Ephrussi, B. *Genetics* **22**, 76–86 (1937).
2. Brenner, S. *Genetics* **77**, 71–94 (1974).
3. von Ehrenstein, G. & Schierenberg, E. in *Nematodes as Biological Models* Vol 1 (ed. Zuckerman, B. M.) 2–71 (Academic, New York, 1980).
4. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *Phil. Trans. R. Soc. B* **275**, 327–348 (1976).
5. Sulston, J. E., Albertson, D. G. & Thomson, J. N. *Dev. Biol.* **78**, 542–576 (1980).
6. Boven, P. *Vidensk. Medd. dansk. naturh. Foren.* **101**, 1–113 (1937).
7. Cassada, R. C. & Russell, R. L. *Dev. Biol.* **46**, 326–342 (1975).
8. Cassada, R. C. & Russell, R. L. *Fedn Proc.* **33**, 1476 (1974).
9. Klass, M. & Hirsh, D. *Nature* **260**, 523–525 (1976).
10. Evans, A. A. F. & Perry, R. N. in *The Organization of Nematodes* (ed. Croll, N. A.) 383–424 (Academic, New York, 1976).

11. Albert, P. S., Brown, S. & Riddle, D. L. *J. comp. Neurol.* (in the press).
12. Swanson, M. M. & Riddle, D. L. *Dev. Biol.* (in the press).
13. Riddle, D. L. *Stadler Genet. Symp.* **9**, 101–120 (1977).
14. Cassada, R. C. in *Developmental Biology: Pattern Formation and Gene Regulation* (eds McMahon, D. M. & Fox, C. F.) 539–547 (Benjamin, Menlo Park, 1975).
15. Lewis, J. A. & Hodgkin, J. A. *J. comp. Neurol.* **172**, 489–510 (1977).
16. Ward, S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 817–821 (1973).
17. Ward, S., Thomson, N., White, J. G. & Brenner, S. *J. comp. Neurol.* **160**, 313–338 (1975).
18. Ware, R. W., Clark, D., Crossland, K. & Russell, R. L. *J. comp. Neurol.* **162**, 71–110 (1975).
19. Herman, R. K. & Horvitz, H. R. in *Nematodes as Biological Models* Vol 1 (ed. Zuckerman, B. M.) 228–261 (Academic, New York, 1980).
20. Wood, W. B., Edgar, R. S., King, J., Lielausis, I. & Henninger, M. *Fedn Proc.* **27**, 1160–1166 (1968).
21. Herman, R. K., Horvitz, H. R. & Riddle, D. L. *Genet. Maps* **1**, 183–193 (1980).

Left-handed DNA in restriction fragments and a recombinant plasmid

Jan Klysik, Steven M. Stirdivant, Jacquelyn E. Larson, Phillip A. Hart & Robert D. Wells

University of Wisconsin, Department of Biochemistry, College of Agricultural and Life Sciences, Madison, Wisconsin 53706, USA
and School of Pharmacy, University of Wisconsin, Madison, Wisconsin 53706, USA

Circular dichroism and ^{31}P -NMR on synthetic oligomers of (dC-dG) inserted within DNA restriction fragments indicate that the right-handed B-structure can exist in close proximity to the left-handed Z-structure. Also, this salt-induced transition to Z-form in a small (dC-dG) segment (1.3%) of a recombinant plasmid markedly influenced the supercoil of the plasmid. These observations have implications for the postulated role of naturally occurring related simple sequences in the regulation of gene activity.

DUPLEX DNA polymers and oligomers containing a strictly repeating dG and dC sequence have been known for some time to display unusual conformational properties¹. High molecular weight $(\text{dG-dC})_n \cdot (\text{dG-dC})_n$ has been synthesized² and characterized by a variety of biochemical and physical approaches (reviewed in refs 1, 3). The presence of high concentrations of salt (NaCl , NaClO_4 , MgCl_2) or ethanol causes this polymer, but not related synthetic polymers, to undergo a marked structural transition as revealed by circular dichroism (CD) and Raman spectroscopy⁴⁻⁷. This cooperative transition has more recently been interpreted (see below) as a conversion of the right-handed helix into a left-handed structure, an observation foreshadowed by the demonstration that $(\text{dI-dC})_n \cdot (\text{dI-dC})_n$ had a very unusual left-handed conformation at normal salt concentrations and could undergo a transition to a typical right-handed DNA-B conformation at high humidities⁸⁻¹⁰.

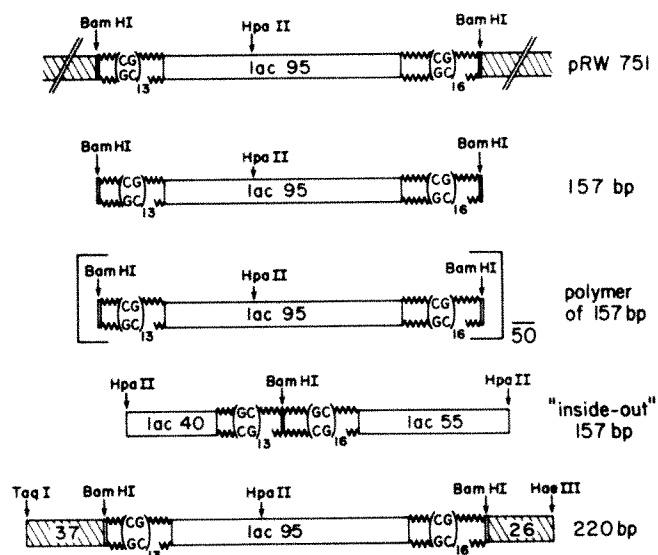
Recent X-ray crystallography studies on short oligomers $[\text{d}(\text{CpGpCpG})]$ and $[\text{d}(\text{CpGpCpGpCpG})]$ revealed the very unusual Z- and Z'-DNA conformations¹¹⁻¹⁵. The Z-conformation differs from the typical right-handed DNA-B structure in that the helix is left-handed with 12 base pairs per turn, the sugar conformations are not all the same but dG is C3'-endo and

dC is C2'-endo and the G residues are in a *syn*, not the orthodox *anti*, conformation. The Z-conformation may correspond to the high salt form of $(\text{dG-dC})_n \cdot (\text{dG-dC})_n$ with the 'inverted' CD spectrum¹¹, particularly as Raman spectroscopy of the crystals used for the X-ray analyses demonstrates the correlation with the high salt form of $(\text{dG-dC})_n$ (A. Rich, personal communication). Also, recent ^{31}P -NMR studies (refs 16, 17 and J. S. Cohen, personal communication) on oligomers with strictly alternating dG and dC residues demonstrates the salt-induced conformational transition from a typical phosphodiester structure (in low salt) to an alternating structure (in high salt). The alternating backbone conformation is presumed to be of the Z-type. It may be that, as proposed^{11,18}, any DNA with a perfectly alternating sequence of purines and pyrimidines adopts a Z-type conformation.

Z-type DNA in a natural genome

The relevance of these observations to natural DNA and their possible implications for the regulation of gene activity is still largely speculative. We therefore decided to construct a recombinant DNA in which poly(dG-dC) was cloned in restriction fragments of plasmid DNA. This approach, based on our

Fig. 1 DNAs used in these studies. pRW751 is a recombinant plasmid derived from pBR322 (ref. 29) which contains a 157-base pair (bp) insert into the 'filled in' *Bam*HI site. The cloning procedure, preparation of plasmid DNA and preparation and characterization of DNA restriction fragments were generally performed as previously described^{1,21,22,24,31}. The 157-base pair fragment was sequenced essentially as described elsewhere²³; the *Bam*HI ends were labelled with ^{32}P and the fragment was cleaved with *Hpa*II. The DNA sequence revealed the presence of the perfectly alternating (dC-dG) sequence shown in this figure along with the presence of the *lac* sequence. The localization of the 26- and 32-base pair (dC-dG) arms relative to the *lac* 95-base pair fragment is precisely known. However, the orientation of the entire 157-base pair insert with respect to the vector is not known. The polymer of the 157-base pair fragment was prepared by a large-scale T4 DNA ligase reaction with the 157-base pair monomer in conditions described previously²² using 100 μg of the 157-base pair fragment. The number average molecular weight (50-mer) of the polymers which resulted from this reaction was established by gel electrophoresis and, more rigorously, by spectrophotometric analysis of the RPC-5 chromatography used in the preparation of the 'inside-out' 157-base pair fragment. This fragment was prepared by *Hpa*II cleavage of the polymer of the 157-base pair fragment. The 'inside-out' molecules are, in fact, a mixture consisting of fragments which are 138, 157 and 176 base pairs long. This mixture was purified from the 'half molecules' (70 and 87 base pairs in length) which were derived from the termini of the polymer. Gel electrophoretic analysis of the purified 'inside-out' 157-base pair molecules showed no detectable 'half molecules' (<1%). The 220-base pair fragment was isolated from pRW751 by a three-step procedure; first, 6 mg of pRW751 were digested with *Hind*III and the resulting products were separated on a preparative polyacrylamide gel. The 780-base pair fragment was extracted from the gel and cleaved with *Taq*I + *Hae*III and the products of this digestion were then fractionated on a preparative polyacrylamide gel. The 220-base pair fragment was extracted from the gel and further purified by column chromatography on DEAE-Sephacel (0.5 $\times$ 0.5 cm) in 10 mM Tris-HCl (pH 7.6), 0.1 mM EDTA. The column was washed with 3 ml of the same buffer and the 220-base pair fragment was eluted with 2.0 M NaCl in 10 mM Tris-HCl (pH 7.6) and 0.1 mM EDTA. The orientation of the 37- and 26-base pair segments of pBR322 on the termini of this molecule relative to the 157-base pair segment was not determined. pBR322 segments are designated with cross-hatching.



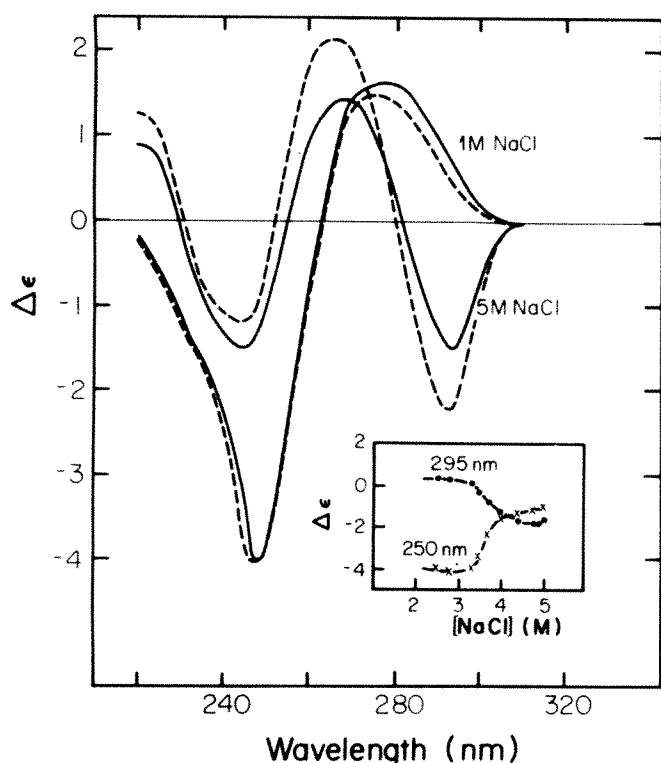


Fig. 2 Circular dichroism spectra of the 157-base pair fragment. The solid lines show the spectra of the 157-base pair fragment in 1.0 and 5.0 M NaCl. The dotted lines show the calculated spectra for a mixture of the *lac* 95-base pair fragment and $(dG-dC)_n \cdot (dG-dC)_n$ at the same molar ratio as in the 157-base pair fragment. The 1.0 and 5.0 M NaCl spectra for the *lac* 95-base pair fragment as well as for the $(dG-dC)_n \cdot (dG-dC)_n$ were taken separately. A solution of $(dG-dC)_n \cdot (dG-dC)_n$ was equilibrated for 4 h at room temperature before the spectra were determined. The extinction coefficients used were the following: *lac* 95-base pair fragment, 6.5×10^3 ; $(dG-dC)_n \cdot (dG-dC)_n$, 7.1×10^3 ; the 157-base pair fragment, 6.7×10^3 . The inset shows the influence of salt concentration on the CD at 295 and 250 nm. The spectra at the different salt concentrations were obtained by diluting the 5.0 M NaCl solution of the 157-base pair fragment with 10 mM cacodylate buffer (pH 7.2) and 0.1 mM EDTA. After each dilution, the sample was equilibrated for 0.5 h at room temperature before the CD spectra were determined. The equilibration time for this fragment was very short; 95% of the transition was obtained after 7 min. $(dG-dC)_n \cdot (dG-dC)_n$ was synthesized as described previously^{2,10}. The *lac* 95-base pair fragment was also prepared and characterized as described previously²⁵ (gift of W. Hillen). All details regarding the CD measurements are described elsewhere²⁵.

methods of cloning^{19,20} and characterizing²¹ other types of DNA polymers with simple sequences, was designed to allow the study of several important questions regarding the Z-structure, including the following: (1) Can Z-DNA exist between flanking segments of B-DNA? (2) What length of (dG-dC) is required to provide a stable Z-conformation? (3) What are the nature and length of the junction between the Z-DNA segments and the flanking B-regions? (4) To what extent does the left-handed Z-type structure influence the biological and physicochemical properties of the neighbouring sequence of random DNA (B-structure) or vice versa? (5) As a Z-type conformation is left handed, does a B $\rightarrow$ Z transition influence the tertiary structure (supercoil) of a plasmid molecule, and will the presence of supercoiling alter the environmental conditions necessary to stabilize a Z-type conformation? (6) Do some biological molecules, such as proteins or small ligands (drugs or carcinogens), specifically interact with the Z-form?

$(dG-dC)_n$ in a plasmid and restriction fragments

Figure 1 illustrates several of the molecules used in this study. A recombinant plasmid (pRW751) was constructed containing an insert 157 base pairs long which was generated by ligating the 95-base pair *lac* operator-promoter fragment²² with a mixture of oligomers prepared by digestion of $(dC-dG)_n \cdot (dC-dG)_n$ with *Fnu*DII. This mixture was ligated into the 'filled-in' *Bam*HI site

of pBR322 by the general procedures described previously^{1,20-22}. The preparation and characterization of this family of recombinant DNAs will be described in detail elsewhere (J. K. and R. D. W., in preparation). A large number of related plasmids were isolated and characterized, many of which were unstable; deletions were observed in the (dC-dG) segments for plasmids which contained (dG-dC) tracts longer than about 40 base pairs (discussed below), but stability was achieved with shorter (dG-dC) tracts. One such plasmid (pRW751) was chosen for further study. It has the structure shown in Fig. 1 and a 157-base pair fragment, excised from it by *Bam*HI and sequenced by methods described previously²³, was shown to consist of the 95-base pair *lac* fragment flanked by strictly alternating dC and dG sequences.

Figure 1 also shows the structure of a polymer of the 157-base pair fragment. This polymer was prepared to internalize the (dC-dG) segments and so evaluate the role of free ends in any possible transition observed in the 157-base pair fragment. The presence of the *Bam*HI sites on the termini of the 157-base pair fragment allowed the formation of this polymer. The polymer was characterized by *Hpa*II digestion followed by quantification of the 'inside-out' 157-base pair fragment relative to the amount of 'half molecules' which are derived from the polymer termini (see below).

Since the first 50-mer of the 157-base pair fragment still contained a small amount of (dC-dG) sequence at the termini, it could not be used to determine rigorously whether the structural transition from a B- to a Z-type structure had to be initiated by (dC-dG) segments at the ends of a DNA molecule or whether it was possible for a segment of (dC-dG) in the interior of a natural sequence to undergo the transition. We therefore produced the inside-out 157-base pair fragment shown in Fig. 1 by *Hpa*II digestion of the 157-base pair 50-mer. This molecule was purified by HPLC on RPC-5 (ref. 24) which cleanly separated the 157-base pair fragment from the small amount of half molecules. Hence, with the inside-out 157-base pair fragment,

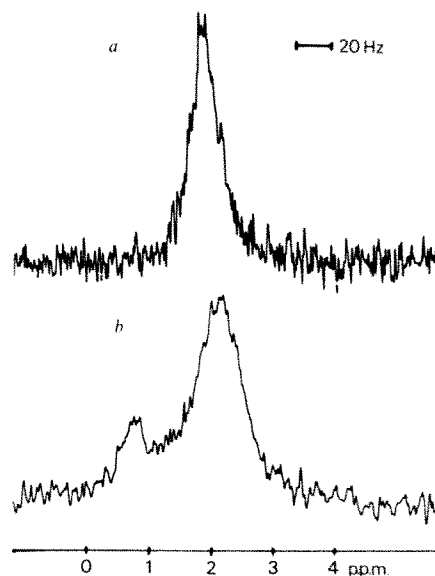


Fig. 3 36.44 MHz ^{31}P -NMR spectra of the 157-base pair fragment in 0.5 and 5.0 M NaCl. The spectra represent a fast Fourier transformation of 4,096 time domain points accumulated following a 90° observing pulse and a 4.2-s delay. Proton resonances were fully and continually decoupled. Exponential multiplication resulting in 0.5 Hz line broadening was applied. Integration was done automatically from manually selected regions of the line integral display. The displayed signals are upfield relative to external 0.95 M H_3PO_4 ; no corrections for salt effects on the deuterium lock signal were made. Temperature was $25 \pm 0.2^\circ\text{C}$. a, 2 mg 157-base pair fragment per ml in D_2O , 0.5 M NaCl, 10 mM cacodylate buffer pH 7.6 and 0.1 mM EDTA; 14,000 accumulations. b, 2 mg 157-base pair fragment per ml in D_2O , 5.0 M NaCl, 10 mM cacodylate buffer pH 7.6 and 0.1 mM EDTA; 19,000 accumulations. The area of the downfield and smaller signal is $\sim 20\%$ of the total resonances. The greater line width at high salt in the 157-base pair fragment could be due to the larger viscosity of the concentrated salt solution, paramagnetic impurities or intermediate phosphorous conformations.

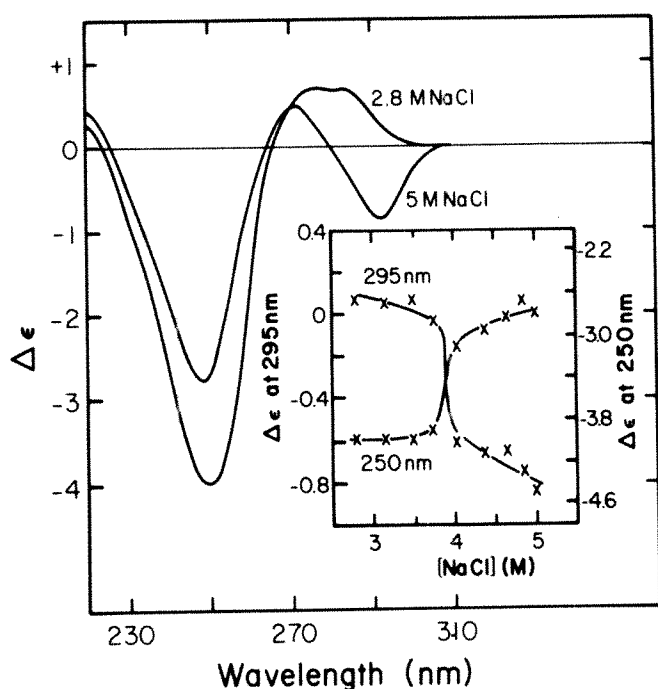


Fig. 4 Circular dichroism spectra of the inside-out 157-base pair fragment in 2.8 and 5.0 M NaCl. The inset shows the CD changes observed at 250 and 295 nm as a function of salt concentration. Other details are described in the legends to Figs 1 and 2.

no (dC-dG) segments are on the termini. It is actually a random mixture of three molecules 138, 157 and 176 base pairs in length which is generated by the *in vitro* ligation used in forming the 50-mer. Note that the (dG-dC) segment in the inside-out 157-base pair fragment contains the *Bam*HI site in its interior.

These three fragments contain identical sequences but in different arrangements. One additional derivative of pRW751 that we have studied is a 220-base pair fragment consisting of the 157-base pair fragment, flanked by segments of pBR322 DNA, 37 and 26 base pairs long (Fig. 1). Hence, the (dC-dG) sequence is flanked by natural sequences which are longer than those for the inside-out 157-base pair fragment. Studies on this 220-base pair fragment supplemented studies with the 50-mer and the inside-out 157-base pair fragment designed to determine if the GATC sequence would adopt a Z-type structure or would remain in a B-conformation in high salt. The 220-base pair fragment differs from the other two DNAs in having (dC-dG) tracts which are uninterrupted by GATC.

Z-type DNA in the 157-base pair fragment

Figure 2 shows the spectra for the 157-base pair fragment in 1.0 and 5.0 M NaCl. In addition, the calculated spectra are shown in identical conditions for the mixture of the appropriate molar ratio of (dC-dG)_n · (dC-dG)_n and the purified 95-base pair fragment. In 1.0 M NaCl, the 157-base pair fragment shows a typical spectrum observed for DNA restriction fragments²⁵; there is very close agreement between the observed spectrum and the spectrum calculated for the component parts. However, in 5.0 M NaCl, the spectrum changes markedly, with negative CD bands at 293 and 245 nm and a positive band at 268 nm. This spectrum agrees well in shape with a calculated spectrum but is lower in amplitude at the peaks at 293 and 268 nm (but not at 245 nm). This general shape in high salt is a result of a marked change in the CD spectrum for (dG-dC)_n · (dG-dC)_n (refs 4-7). The discrepancy between the observed CD curve at 5.0 M NaCl and the calculated curve indicates that a portion of the (dC-dG) arms does not undergo the salt-induced transition to the Z-type conformation (discussed below). Also, we conclude that most (perhaps all) of the 95-base pair *lac* segment remains in a B-structure.

The inset in Fig. 2 shows the cooperative nature of the change in CD at 295 and 250 nm as expected for the salt-induced B → Z

change in conformation in the (dC-dG) arms. In contrast, no cooperative change is observed in similar conditions for the 95-base pair *lac* fragment alone. The midpoint of this transition (~3.6 M) is substantially higher than for (dG-dC)_n · (dG-dC)_n (~2.5 M)⁴, presumably due to the junction with the *lac* sequence.

³¹P-NMR studies were performed to confirm that the inversion in the CD spectrum was due to formation of a Z-type conformation in the (dC-dG) ends of the 157-base pair fragment. Figure 3 shows the spectra observed in 0.5 and 5.0 M NaCl. A single symmetrical peak is observed in the lower salt whereas a second resonance, which is 1.45 p.p.m. downfield from the main resonance, was observed in 5.0 M salt. This result is clearly consistent with previous observations^{16,17}, which documented a splitting of similar magnitude in the phosphorous spectrum of the high salt form of (dG-dC)_n · (dG-dC)_n oligomers.

Integration of the peaks in Fig. 3b indicates that approximately 20% of the phosphorous resonances are in the second small peak which was shifted 1.45 p.p.m. As the (dG-dC) segment comprises ~37% of the 157-base pair fragment, this suggests that virtually all the (dC-dG) tracts have converted to a Z-type conformation. The finding that the NMR studies indicate transition of a larger percentage of the (dC-dG) sequences than is deduced from the CD measurements, may be due to the fact that different conformational features of the molecule are being investigated.

Laser Raman spectroscopy studies (in collaboration with R. M. Wartell) are in complete agreement with these results.

Z-type DNA between flanking B-DNA segments

Figure 4 shows the spectra observed in 2.8 and 5.0 M NaCl for the inside-out 157-base pair fragment. This molecule is of interest because the (dG-dC) segments are totally embodied within the 157-base pair fragment; there are no (dG-dC) segments at the termini.

A typical DNA spectrum²⁵ is observed for this fragment in 2.8 M NaCl. However, in 5.0 M salt the spectrum changes substantially. The trend of the changes is in agreement with the changes found for the 157-base pair fragment (Fig. 2) but their magnitude is less than observed for the 157-base pair fragment and less than calculated for the mixture of the 95-base pair *lac* fragment and the (dC-dG) arms. Thus, the two junctions of the (dG-dC) segments (at the *lac* fragment junctions and at the *Bam*HI site) clearly diminish the capacity of the (dG-dC) segment to undergo the B → Z conformational change. The inset in Fig. 4 shows that the observed transitions are cooperative, as expected for the B → Z conformational transition.

Two other fragments were studied which contained (dC-dG) segments between two naturally occurring B-DNA segments. The polymer containing an average of 50 tandem copies of the 157-base pair fragment (Fig. 1) was also studied by CD. This fragment differs from the inside-out 157-base pair molecule in length but not sequence. The CD spectra of this polymer at 2.8 and 5.0 M NaCl (data not shown) are almost identical to those shown in Fig. 4 for the inside-out 157-base pair fragment. Also the transition is cooperative, with a midpoint which is identical to that found for the inside-out 157-base pair fragment (3.9-4.0 M). Thus, the length of the molecule has no apparent influence on these properties.

Moreover, the CD properties of the 220-base pair fragment (Fig. 1) show the same trends as observed for the inside-out 157-base pair fragment and for the polymer of the 157-base pair fragment, but the magnitude of the change is somewhat less, as expected from its smaller (dC-dG) content. The quantitation of these transitions and their interpretation regarding DNA structure are described below. Because the same type of CD transition is observed with the 220-base pair fragment, we conclude that the transition is not a function of some peculiarity of the other three fragments (157-base pair, polymer of 157-base pair,

and inside-out 157-base pair fragments) but is a general property of the (dC-dG) sequence embodied in a natural DNA molecule.

Junction between B- and Z-type conformations

It is possible to calculate the fraction of the (dC-dG) segments which are in a Z-type conformation for all the fragments studied (Fig. 1), as virtually no CD signal at 295 nm is observed in 5.0 M NaCl for the *lac* 95-base pair fragment. Hence, any negative signal observed for the fragments at 295 nm must be due to (dC-dG) segments which are transformed into the Z-type conformation. Thus, by comparison with the calculated CD signal for the component parts, we can determine the amount of Z-type structure. For these calculations, we shall use an average of 29 base pairs for the (dC-dG) tracts (average of 26 and 32). For the 157-base pair fragments, we deduce that 65% (or 19 out of an average of 29 base pairs in each segment) are in a Z-type conformation. As each (dC-dG) segment is joined only once to the *lac* 95-base pair segment, 10 base pairs are in a conformation other than Z-type. We believe that these are at the junction of the (dC-dG) segment and the *lac* 95-base pair fragment.

For the inside-out 157-base pair fragment, 36% (or 10 base pairs out of an average length of 29) are in a Z-type conformation. As the (dG-dC) segment interfaces at both ends with a random type sequence, $19/2 = 9.5$ base pairs (or 10 base pairs) constitute the junction region. This, of course, presumes that the GATC *Bam*HI site disrupts the Z-type conformation of the (dG-dC) tract.

For the polymer of the 157-base pair fragment, 14% (or 4 base pairs out of an average of 29) are in a Z-type conformation. As the (dC-dG) region in this polymer also interfaces at both ends with random DNA sequences, $25/2 = 12.5$ base pairs constitute each of the junction regions which are not in a Z-type conformation. Similarly, for the 220-base pair fragment, 16% (5 base pairs out of an average of 29) are in a Z-type conformation, leaving a junction region of 12 base pairs.

Thus, the length of the junction regions for all four molecules ranges from 9.5 to 12.5 base pairs. The remarkable consistency in the length of this region implies the necessity for a transition region of ~ 11 base pairs from a B-conformation to the Z-type structure. The structure of this intermediate region (approximately one turn of helix) is unknown. Wing *et al.*²⁶ have determined the crystal structure of d(CGCGAATTCGCG); their finding that the CGCG termini are not in a Z-structure is consistent with our results.

These calculations make several assumptions, including the following: (1) The CD signal from the segments of *lac* and pBR322 DNA in these restriction fragments is the same as for the 95-base pair segment alone. That is, the conformation of the natural sequences as judged by CD is not influenced by their junction to Z-type DNA. (2) The GATC which comprises the *Bam*HI site is not in a Z-type conformation when the surrounding (dC-dG) tracts are in a Z-type structure for the cases of the polymer of the 157 base pairs and the inside-out 157-base pair fragment. (3) The junction regions (9.5–12.5 base pairs in length) of (dC-dG) residues which are not in a Z-type conformation give no CD signal at 295 nm.

Z-DNA in a supercoiled plasmid

It was of interest to determine if a Z-type conformation can exist within a covalently closed and circular DNA. On the basis of previous studies on supercoiled DNA, it is logical that a B $\rightarrow$ Z transition in a (dC-dG) segment should alter the number of supercoil turns in a plasmid DNA because the supercoil number is dictated by the specific linking difference which is sensitive to conformational changes in the primary helix (reviewed in refs 27, 28).

pRW751 was a suitable molecule for studying this question. The total length of (dC-dG) sequences in pRW751 is 58 base

pairs or 1.3% of the length of the total recombinant plasmid (4,516 base pairs). The unwinding of one turn of primary helix removes one right-hand supercoil. In addition, each turn of left-handed primary helix which is induced will also remove one right-hand supercoil. Thus, we can calculate that a complete transition of 58 base pairs of (dC-dG) from B $\rightarrow$ Z would theoretically remove 10.3 supercoils from pRW751 ($58/10.4$ base pairs per turn + $58/12$ base pairs per turn = 10.3).

We have established conditions for running agarose gels in NaCl solutions up to 5.0 M (Fig. 5) to determine the number of supercoil turns in a DNA. pRW451 served as a control for these electrophoresis experiments as it is identical to pRW751 apart from containing a 174-base pair insert (formerly a *Hha*I fragment of pBR322 DNA which contains 54% G + C) instead of the 157-base pair insert in pRW751. This 174-base pair insert does not contain unusual sequence features or stretches of (dC-dG) sequence²⁹.

Figure 5 shows the gel electrophoresis patterns observed for pRW451 and pRW751 in 3.0 and 4.5 M NaCl. In 3.0 M, the partially relaxed family of topoisomers for both DNAs comigrate (lanes a and b). For both DNAs in these electrophoresis conditions, the family ranges from approximately five right-

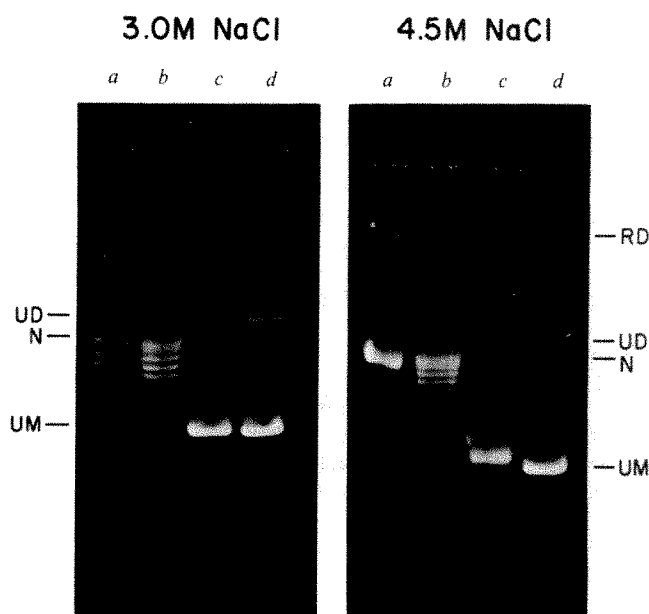


Fig. 5 High salt gels on supercoiled plasmids. Monomeric supercoiled pRW751 was separated from higher forms (such as dimer, tetramer) by HPLC on RPC-5 using a 1.5 mm $\times$ 20-cm column at 43 °C. Elution was with a 0.65–0.8 M KCl gradient using the general techniques described previously²⁴. pRW451 contains the 174-base pair *Hha*I fragment from pBR322 cloned into the filled-in *Bam*HI site of pBR322 (G. D. Staffeld and R. D. W., unpublished). pRW451 was isolated in a similar manner. Partially relaxed pRW751 and pRW451 were prepared by treatment with wheat-germ topoisomerase (gift of R. R. Burgess). Reaction mixtures (100 μ l) contained 20 μ g DNA and 8 μ g enzyme in 50 mM Tris-HCl pH 7.6, 0.1 mM dithiothreitol, 10 mM MgCl₂ and 2.5 mM spermidine. The reaction was for 30 min at 37 °C which gives a terminal reaction product of right-handed supercoiled topoisomers. The unrelaxed and partially relaxed plasmids were electrophoresed through 1% agarose gels with a buffer of 80 mM Tris-HCl, 40 mM sodium acetate, 4.4 mM disodium EDTA and 0.525 M acetic acid, all at pH 8.3, plus either 3.0 or 4.5 M NaCl. The gels were 8 cm $\times$ 14 cm $\times$ 1 mm. The DNA samples were equilibrated in a loading buffer identical to the electrophoresis buffer but with an identical molarity of NaClO₄ substituted for NaCl and 0.05% bromophenol blue added. The gels were electrophoresed for 36 h at 25 V for the 4.5 M NaCl gel and 20 V for the 3 M NaCl gel. The gels were desalted by soaking in water for 10 min and then stained with 100 μ g ml⁻¹ ethidium bromide. Abbreviations are: UM, unrelaxed monomer; N, nicked monomer; UD, unrelaxed dimer; RD, relaxed dimer. The following samples were electrophoresed in the designated lanes: a, partially relaxed pRW751; b, partially relaxed pRW451; c, unrelaxed pRW751; d, unrelaxed pRW451. Resolution of fully relaxed plasmid from supercoiled DNA containing +1 or -1 supercoils is difficult in these conditions. Similarly, the logarithmic nature of topoisomer migration prevents the resolution of individual bands for topoisomers containing ~ 10 supercoils.

handed supercoil turns to fully relaxed DNA (the exact number of supercoil twists is uncertain due to our inability to resolve topoisomers near the fully relaxed position, but we believe that this number is accurate within at least ± 1). This experiment was performed with plasmids which were partially relaxed with topoisomerase to obtain better resolution of the family of topoisomers and thus enhance our ability to detect smaller changes in supercoil number. Figure 5, lanes *c* and *d*, shows that the unrelaxed plasmids co-migrate in 3.0 M NaCl.

When similar studies were done in 4.5 M NaCl, no change was observed for the migration pattern of the partially relaxed control plasmid (pRW451) (Fig. 5, lane *b*). However, Fig. 5, lane *a*, shows that all the topoisomers for pRW751 co-migrated in these conditions at the fully relaxed position. Thus, a loss of as many as five supercoil turns was observed, but there was no apparent generation of left-handed supercoil twists.

This striking result clearly demonstrates that a salt-induced structural change occurred in pRW751 which had a dramatic influence on the supercoil number, and is consistent with a conformational transition from B- to a Z-type structure for at least a portion of the (dC-dG) tracts in pRW751.

A small difference was observed in the migration of pRW451 and pRW751 at 4.5 M NaCl (Fig. 5, lanes *c*, *d*). Based on our experience with topoisomer migration patterns, we believe this difference is equivalent to a loss of four to eight supercoil turns. However, the relatively poor resolution of topoisomers of high superhelical density makes precise quantification difficult, especially in these electrophoresis conditions.

Electrophoresis studies were performed on the partially relaxed plasmids at several salt concentrations between 3.5 and 4.0 M in addition to those shown in Fig. 5 to determine the degree of cooperativity. At salt concentrations of 3.5 M to at least 4.0 M, the topoisomers with more supercoil turns decreased in supercoil number at lower salt concentrations than topoisomers with fewer supercoil turns, suggesting that supercoiling contributes a favourable free energy to the B- to Z-type transition.

Thus, the studies with both the partially relaxed and unrelaxed plasmids indicate an unwinding of about five supercoil turns (conceivably not fewer than four and not more than eight) in 4.5 M NaCl. In the calculations described above, if the entire (dC-dG) segment in pRW751 converted from a B $\rightarrow$ Z conformation, we should have observed a loss of 10.3 supercoil turns. As this large an effect was not observed, the entire length of the (dC-dG) segments may not be in a Z-type structure, as predicted from the CD results.

Conclusions

Circular dichroism and ^{31}P -NMR studies on fragments show that the (dC-dG) segments undergo a cooperative transition to a left-handed conformation in the presence of high salt. The natural sequences from the *lac* operon and pBR322 do not appear to convert to a left-handed conformation. In all cases a junction of ~ 11 base pairs seems to exist between the left-handed conformation and B-DNA which is in neither a Z-type nor a B-conformation but must be intermediate in nature.

Moreover, electrophoretic analyses of plasmids containing the (dG-dC) sequence demonstrate a cooperative transition from a right-handed conformation to a left-handed Z-type structure. The presence of the (dG-dC) segment in the amount of 1.3% of the recombinant plasmid is sufficient to have a profound influence on the supercoil properties of the DNA. These observations are consistent with prior determinations on the extreme flexibility of the DNA chain (see ref. 30, reviewed in ref. 1).

Recombinant plasmids related to pRW751 commonly showed deletions in the (dG-dC) tracts, a type of behaviour unprecedented in our experience with a substantial variety of cloning studies^{1,19-22,24,31-33}. It seems that (dG-dC) segments longer than ~ 40 base pairs may not be stable in *Escherichia coli* (J. K. and R. D. W., unpublished).

Biological implications of left-handed DNA

The demonstration that different segments of one DNA chain can co-exist with conformations as different as a right-handed B-helix and a left-handed Z-type structure may have profound implications for the role of DNA structure in regulatory processes. Whereas this is not the first implication of DNA conformations and properties as important components in gene regulation or of left-handed DNA structures (reviewed in ref. 1), it is the first clear demonstration of left-handed DNA which is adopted by a naturally replicating molecule and by portions of restriction fragments isolated from that plasmid.

A segment of left-handed DNA in a viral or cellular DNA which is principally comprised of right-handed conformations would be an easily recognizable locus for binding of specific proteins or ligands³⁴ due to the dramatic difference in conformation. Moreover, the short region (approximately one turn of helix) at the interface between the B- and Z-conformations probably consists of intermediate structures which would also provide unique interaction contacts for ligand binding.

Our studies demonstrate that changes in the conformation of the primary helix in as little as 1.3% of pRW751 can profoundly influence its supercoil structure. Thus, conformational transitions in tiny portions of a chromosome can have large effects on the overall topography of a genome. As the extent of supercoiling has differential effects on the transcription properties of different genes (reviewed in ref. 1), this implicates conformational transitions in the primary helix as a genetic regulatory mechanism.

Our results clearly demonstrate that the cooperative salt-induced transition can be initiated within a polynucleotide chain. It is possible that a smooth and nondisruptive transition may be sterically and energetically feasible between B- and Z-conformations, as shown³⁵⁻³⁷ for the interface of B- and A-structures.

It is uncertain if appropriate natural sequences, in fact, undergo a transition to a Z-structure *in vivo*. However, several microorganisms (for example, halobacteria) as well as crustacea have high intracellular salt concentrations. Moreover, agents other than high salt may effect the B $\rightarrow$ Z transition *in vivo*.

Do sequences exist in viral or cellular DNAs which can adopt a left-handed conformation? We are unaware that a long tract of perfect (dG-dC) has yet been found in a duplex DNA. However, 7-base pair lengths of this sequence are found in a hairpin region of four related parvovirus genomes (reviewed in ref. 1). Moreover, a number of related repeating dinucleotide sequences have been found³⁸⁻⁴³.

This work was supported by NIH grant CA 20279 and NSF grants PCM 77-15033 and PCM 77-19927.

Received 3 November 1980; accepted 2 February 1981.

- Wells, R. D. *et al.* *Prog. Nucleic Acid Res. molec. Biol.* **24**, 167-267 (1980).
- Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E. & Cantor, C. R. *J. molec. Biol.* **54**, 465-497 (1970).
- Wells, R. D. *et al.* *Crit. Rev. Biochem.* **4**, 305-340 (1977).
- Pohl, F. M. & Jovin, T. M. *J. molec. Biol.* **67**, 375-396 (1972).
- Pohl, F. M., Jovin, T. M., Baehr, W. & Holbrook, J. J. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3805-3809 (1972).
- Pohl, F. M., Ranade, A. & Stockburger, M. *Biochim. biophys. Acta* **335**, 85-92 (1973).
- Pohl, F. M. *Nature* **260**, 365-366 (1976).
- Grant, R. C., Harwood, S. J. & Wells, R. D. *J. Am. chem. Soc.* **90**, 4474-4476 (1968).
- Mitsui, Y. *et al.* *Nature* **228**, 1166-1169 (1970).
- Grant, R. C., Kodama, M. & Wells, R. D. *Biochemistry* **11**, 805-815 (1972).
- Wang, A. H.-J. *et al.* *Nature* **282**, 680-686 (1979).
- Crawford, J. L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 4016-4020 (1980).
- Wang, A. H.-J. *et al.* *Science* **211**, 171-176 (1980).
- Crawford, J. L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Drew, H., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. *Nature* **286**, 567-573 (1980).
- Patel, D. J., Canuel, L. S. & Pohl, F. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2508-2511 (1979).
- Simpson, R. T. & Shindo, H. *Nucleic Acids Res.* **8**, 2093-2103 (1980).
- Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W. & Ratliff, R. L. *Nature* **283**, 743-745 (1980).
- Klein, R. D., Selsing, E. & Wells, R. D. *XIII int. Congr. Biochem.* 01-1-H26 (1979).
- Selsing, E. & Wells, R. D. *Nucleic Acids Res.* **6**, 3025-3040 (1979).
- Klein, R. D., Selsing, E. & Wells, R. D. *Plasmid* **3**, 88-91 (1980).
- Hardies, S. C. *et al.* *J. biol. Chem.* **254**, 5527-5534 (1979).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Wells, R. D. *et al.* *Meth. Enzym.* **65**, 327-347 (1980).
- Hillen, W. & Wells, R. D. *Nucleic Acids Res.* **8**, 5427-5444 (1980).

26. Wing, R. *et al.* *Nature* **287**, 755–758 (1980).
27. Bauer, W. R. A. *Rev. Biophys. Bioengng* **7**, 287–313 (1978).
28. Wang, J. C. *Trends biochem. Sci.* **5**, 219–221 (1980).
29. Sutcliffe, J. G. *Cold Spring Harb. Symp. quant. Biol.* **43**, 77–90 (1978).
30. Wartell, R. M., Larson, J. E. & Wells, R. D. *J. biol. Chem.* **250**, 2698–2702 (1975).
31. Hardies, S. C. & Wells, R. D. *Gene* **7**, 1–14 (1979).
32. Panayotatos, N. & Wells, R. D. *J. molec. Biol.* **135**, 91–109 (1979); *Nature* **280**, 35–39 (1979).
33. Müller, U. R. & Wells, R. D. *J. molec. Biol.* **141**, 1–24 (1980); *J. molec. Biol.* **141**, 25–41 (1980).
34. Sage, E. & Leng, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4597–4601 (1980).
35. Selsing, E., Wells, R. D., Early, T. A. & Kearns, D. R. *Nature* **275**, 249–250 (1978).
36. Selsing, E. & Wells, R. D. *J. biol. Chem.* **254**, 5410–5416 (1979).
37. Selsing, E., Wells, R. D., Alden, C. J. & Arnott, S. J. *J. biol. Chem.* **245**, 5417–5422 (1979).
38. Nishioka, Y. & Leder, P. *J. biol. Chem.* **255**, 3691–3694 (1980).
39. Sures, I., Lowry, J. & Kedes, L. H. *Cell* **15**, 1033–1044 (1978).
40. Maio, J. J. *Handbk Biochem.* **2**, 379–404 (1976).
41. Kimmel, A. R. & Firtel, R. A. *Cell* **16**, 787–796 (1979).
42. Schachman, H. K., Adler, J., Radding, C. M., Lehman, I. R. & Kornberg, A. *J. biol. Chem.* **235**, 3242–3249 (1960).
43. Burd, J. F. & Wells, R. D. *J. molec. Biol.* **53**, 435–459 (1970).

Evolution of the 87A and 87C heat-shock loci in *Drosophila*

A. J. Leigh Brown & D. Ish-Horowicz

Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, London NW7 1AD, UK

Several species related to Drosophila melanogaster have two loci containing sequences which code for the major heat-shock protein of molecular weight 70,000 (hsp70). These sequences are arranged as two pairs of inverted repeats. The structure of the 87C locus in D. melanogaster seems to have arisen through DNA insertion. Hsp70-coding sequences at the two loci are not evolving independently. Gene conversion between hsp70 genes has apparently occurred both within and between loci.

In higher eukaryotes, a single enzymatic function is frequently encoded at more than one locus. The duplicated loci need not be closely linked but are likely to derive from a single ancestral locus. A second form of gene reiteration is the presence of tandem arrays of repeated genes, for example ribosomal DNA and 5S genes. We have studied the evolution in *Drosophila* of a gene system showing both styles of duplication for which detailed molecular analysis is possible. Several features have emerged which suggest that classical views on mutation and evolution in such systems are inadequate.

Exposure of *Drosophila* tissue to a temperature of 37 °C activates a characteristic set of about nine chromosomal puffs and also induces the synthesis of a new set of about nine proteins (reviewed in ref. 1). As the heat shock inhibits previous RNA and protein synthesis, the resulting pattern of macromolecular synthesis is greatly simplified. Several laboratories have used heat-shock mRNA to select cloned genomic DNA fragments that encode the major 70,000-molecular weight heat-shock protein (hsp70). Such DNA fragments have been localized to two heat-induced puffs (87A and 87C)^{2–4}. This has been confirmed by genetic analysis using mutations that delete one or other locus. In this way, the two hsp70-coding loci were mapped to the single bands 87A7 and 87C1 (ref. 5). The sibling species

Drosophila simulans also makes a 70,000-molecular weight heat-induced protein⁶, and *in situ* hybridization shows that it too is encoded at two loci, 87A and 87C (ref. 2). A second type of heat-induced sequence, called 'αβ' sequences, is found at 87C1 but not at 87A7 in *D. melanogaster*⁷. These sequences seem to be absent at both hsp70 loci in *D. simulans*, although, like *D. melanogaster*, they are found at the chromocentre².

The arrangement of hsp70-coding sequences in *D. melanogaster* has been elucidated by use of cloned DNA sequences. The total number of hsp70 genes varies from five to about eight according to genotype, there being two copies at 87A7 and three to five at 87C1 (refs 8–10). (In this discussion, 'gene' refers to an mRNA-coding element and adjacent conserved sequences. 'Locus' designates the cytogenetic origin, that is, 87A7 or 87C1.) Each of the hsp70 genes seems to be organized within a 2.5-kilobase conserved element (z), comprising the 2.1-kilobase mRNA-coding region (z_c) and the 400 5'-terminal bases that are not transcribed (z_{nc})^{3,8,9,11}. Nevertheless, the z elements at 87A7 and 87C1 differ by some characteristic restriction sites^{8,10}.

Studies on gene arrangement at the hsp70 loci have shown that the two hsp70 genes at 87A7 are about 1.7 kilobases apart (Fig. 1) in an opposite orientation^{9,12,13}. The copies at 87C1 are

Table 1 Comparison of 87A and 87C heat-shock loci in *D. melanogaster*, *D. simulans* and *D. mauritiana*

	Total no. of restriction sites used in comparison	Sites	Restriction sites in common and nucleotide substitution per site (δ)			
			With <i>D. melanogaster</i>		With <i>D. simulans</i>	
			δ	Var (δ)	δ	Var (δ)
<i>D. melanogaster</i>						
Protein coding	26					
Non-coding	39 (<i>D. mauritiana</i>)					
	37 (<i>D. simulans</i>)					
<i>D. simulans</i>						
Protein coding	24	22	0.0213	0.000152		
Non-coding	38 (<i>D. mauritiana</i>)	31	0.0295	0.000145		
	37 (<i>D. melanogaster</i>)					
<i>D. mauritiana</i>						
Protein coding	24	22	0.213	0.000152	24	0.00
Non-coding	34 (<i>D. melanogaster</i>)	30	0.327	0.000165	30	0.0257
	32 (<i>D. simulans</i>)					0.000132

Division of the hsp70 genes into protein-coding and non-protein-coding regions was made from ref. 35. The protein-coding region includes the 3'-terminal *SalI* sites and the 5' *BglII* sites. It does not include the 5' *PstI*, *SalI* or *EcoRI* (*D. mauritiana*) sites. The total number of restriction sites is the number of sites in one species which could have been detected in a second. This number varies in different comparisons because only the restriction sites closest to the z elements could be mapped. The mean frequency of nucleotide substitution and its variance (Var δ) were estimated after Nei and Li⁴².

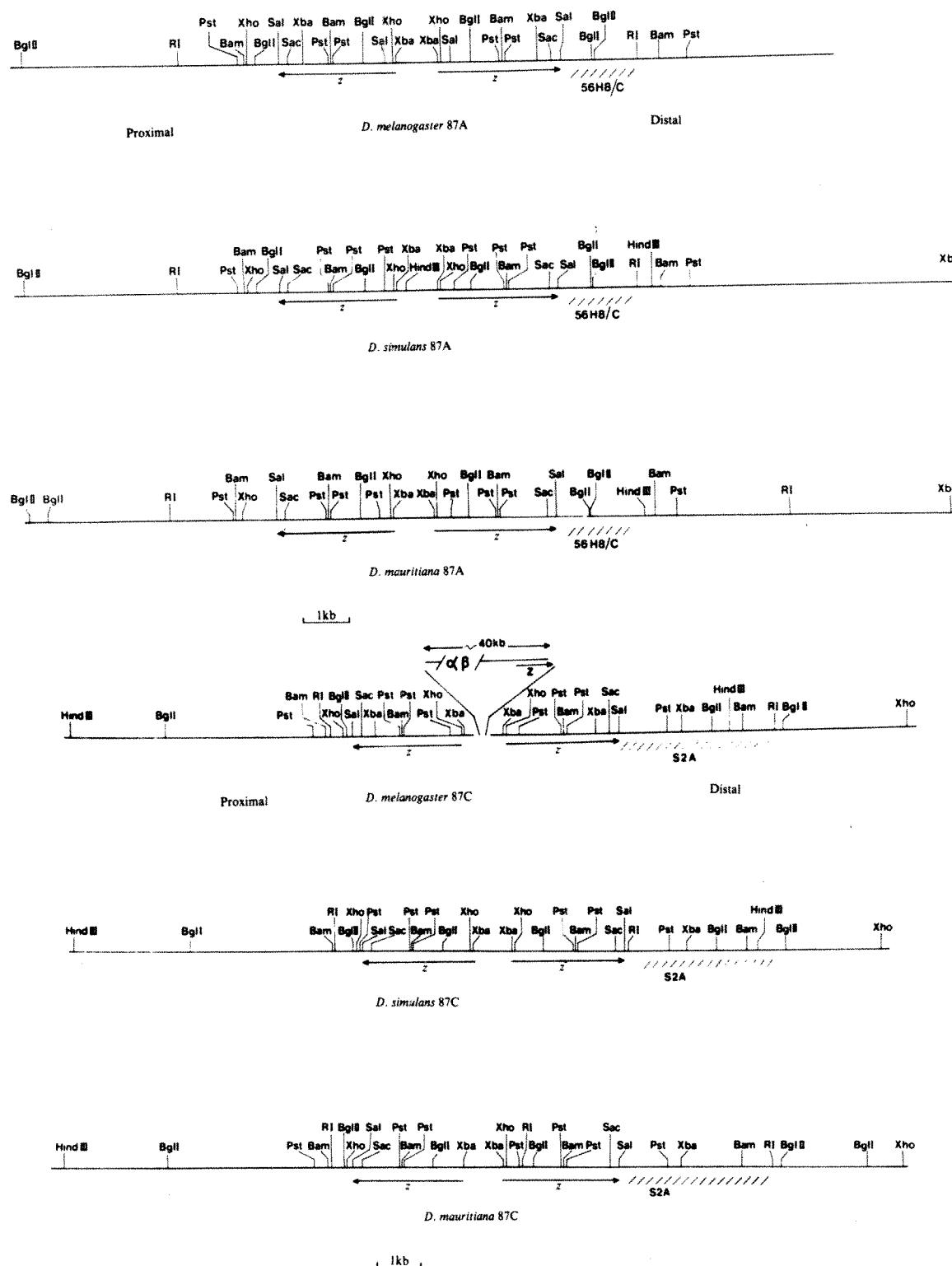


Fig. 1 Organization of hsp70-coding loci. DNA extraction, purification, digestion and electrophoresis were as described previously^{8,13}. DNA was prepared from adult flies of inbred strains of *D. simulans* and *D. mauritiana* obtained by single pair matings for several generations. The DNA was analysed by the technique of Southern¹⁵ as adapted by Wahl *et al.*⁴³. The organization of hsp70-coding loci in *D. melanogaster* is redrawn from ref. 13. (The region depicted in the inset at the 87C1 locus is not drawn to scale.) The hsp70-coding sequences in our inbred strains of *D. simulans* and *D. mauritiana* were positioned and oriented at the cytogenetic loci using two locus-specific probes. Plasmid 56H8/C is derived from plasmid 56H8 (ref. 16) and contains a subcloned *SalI*-*EcoRI* fragment that lies just downstream of the distal hsp70 gene at 87A7. This probe hybridizes only to restriction fragments of *D. melanogaster* known to come from 87A7 (data not shown). It allows us to visualize restriction fragments from 87A7 and to orientate the 87A7 restriction maps on the chromosome. Similarly, plasmid S2A is derived from a cosmid clone cHS2 bearing two tandem genes from the distal region at 87C1 in *D. melanogaster* (A. Udvardy, personal communication). We have subcloned the *SalI*-*EcoRI* fragment immediately downstream from the most distal *z* element of cHS2, making a probe specific for 87C1. The *z* elements whose 3'-adjacent DNA sequences hybridize to these probes are aligned in the figure. Note that the designation of the homologous genes as 'proximal' and 'distal' is inappropriate in *D. simulans* and *D. mauritiana* because of the large inversion on this arm of the third chromosome which is found in these species¹⁴. To avoid confusion only the *D. melanogaster* restriction maps have been thus labelled. kb, Kilobase.

present in two domains, a single proximal sequence separated from tandem distal sequences by about 38 kilobases containing the $\alpha\beta$ heat-induced sequences¹³ (Fig. 1).

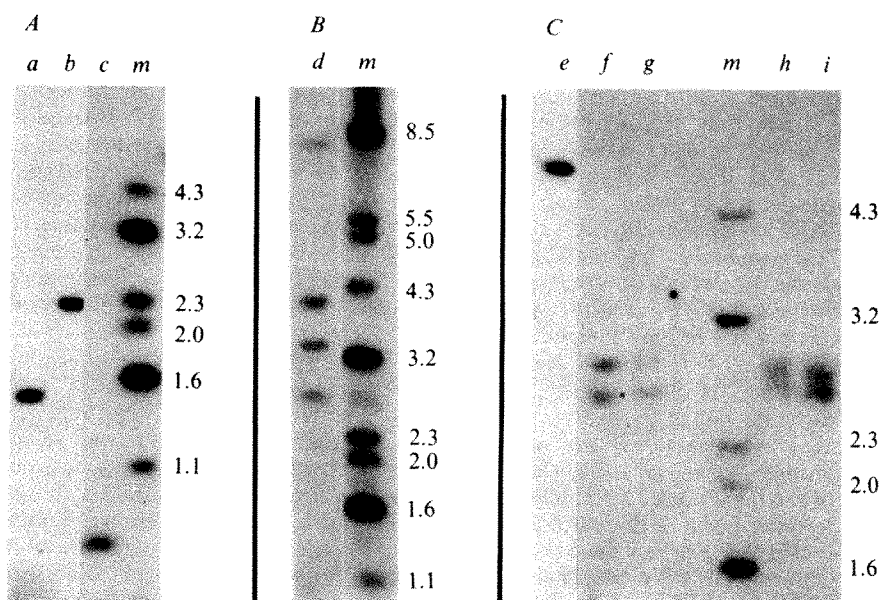
In this article we compare the gene organization of hsp70 sequences in three closely related species, *D. melanogaster*, *Drosophila mauritiana* and *D. simulans*, which show very similar polytene chromosome banding patterns¹⁴. We deduce the probable ancestral organization, and use the restriction maps to estimate the frequency of nucleotide substitution between the species. We also find that the hsp70 genes within a species are much more alike than would be expected if they were evolving independently. This suggests that there is gene conversion both within and between hsp70 loci and we discuss a possible mechanism by which this could occur.

Organization of hsp70-coding sequences in *D. simulans* and *D. mauritiana*

Hsp70 genes of *D. simulans* and *D. mauritiana* are found at two loci, 87A and 87C, as in *D. melanogaster* (ref. 2; S. M. Pinchin, personal communication). We have studied their organization in adult fly DNA using the transfer hybridization method of Southern¹⁵. Figure 1 shows the restriction maps of the hsp70 loci of *D. simulans* and *D. mauritiana*. The maps reveal that the hsp70-coding regions lie within a conserved z element, as in *D. melanogaster*. In particular, the *Xba*I and *Xho*I restriction sites characteristic of z_{nc} are conserved, strengthening the idea that z_{nc} is important for hsp70 expression³. In Fig. 2 we show certain restriction digests that establish the conservation of z elements and others that demonstrate the presence of four coding sequences in these species. Our *D. simulans* and *D. mauritiana* stocks did not show the variation in hsp70 gene number that is found in *D. melanogaster*, but we have not examined other independent lines.

Both hsp70 loci of *D. simulans* and *D. mauritiana* have the same organization—two hsp70 genes in opposite orientations (Fig. 1). We have demonstrated the presence of inverted-repeat hsp70 genes by the presence of S_1 nuclease-resistant genes in denatured genomic DNA (Fig. 2C). We find two S_1 -resistant ('snapback') bands of 2.6 and 2.8 kilobases which include the whole z element. [The absence of an *Xho*I site in an 87C gene of *D. mauritiana* is due to a restriction site polymorphism (A.J.L.B., unpublished data).] This orientation is confirmed by digestion of chromosomal DNA with the restriction enzyme *Sac*I which cuts every hsp70 gene near the 3' end^{4,13}. Only a single 5.3-kilobase band is found (Fig. 2C), showing that the genes are paired and the same distance apart.

Fig. 2 A, Hsp70 genes are arranged in a z element in *D. mauritiana* and *D. simulans*. Autoradiograms of transferred agarose gels were hybridized with the plasmids pPW232.1 and pPW229.1, probes for the 5' and 3' halves of the hsp70 gene respectively². In lanes *m* are molecular-weight markers of size indicated in kilobases. In lanes *a*–*c*, *D. mauritiana* DNA has been digested with *a*, *Xba*I/*Bam*HI (pPW232.1); *b*, *Xba*I/*Sac*I (pPW232.1); *c*, *Bam*HI/*Sac*I (pPW229.1). B, There are four copies of hsp70-coding sequences in *D. mauritiana* and *D. simulans*. *d*, An *Xba*I/*Bgl*II digest of *D. mauritiana* DNA probed with pPW229.1. *D. simulans* gives identical patterns with all the above digests. C, 'Snapback' hsp70 genes in species related to *D. melanogaster*. Lane *e* contains a *Sac*I digest of *D. mauritiana* DNA (an identical pattern is obtained with *D. simulans*). In lanes *f*–*i*, DNA was denatured, reannealed and digested with S_1 nuclease exactly as described previously¹³, before electrophoresis and transfer hybridization to plasmid pPW229.1. Lane *f* is *D. teisseri*, *g* is *D. yakuba*, *h* is *D. simulans*, and *i* is *D. mauritiana*.



Evolution of the organization of hsp70 genes in *Drosophila*

The members of the *D. melanogaster* subgroup are closely related. A previous study on the polytene chromosome banding patterns has suggested that *D. simulans* and *D. mauritiana* are more closely related to each other than to *D. melanogaster*¹⁴. Two more distantly related species, *Drosophila teisseri* and *Drosophila yakuba*, also have two pairs of hsp70 genes both arranged as inverted repeats (Figure 2C). This indicates that this was the ancestral organization which existed at both 87A7 and 87C1 before the separation of *D. melanogaster* from the other species.

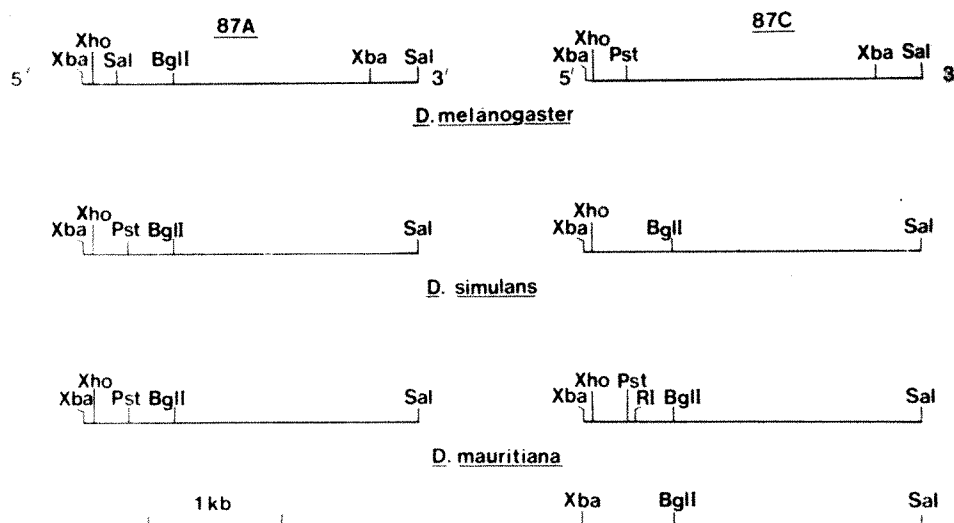
The 87C1 locus in *D. melanogaster* is very different from all the other hsp70 loci (Fig. 1)¹³. It must have arisen through DNA insertion between the two diverging hsp70 genes of the ancestral locus. The intervening region between hsp70 domains at this locus is very large and contains many repeated sequences, including the α and β elements which are also found at the chromocentre and some other sites⁷. The α and β elements are also present at the chromocentre in *D. simulans*² and so may be presumed to have been present before the speciation of *D. simulans* and *D. melanogaster*. As the chromocentral α and β elements of *D. melanogaster* are non-adjacent¹⁶, it would require at least two events to introduce contiguous $\alpha\beta$ elements at 87C1.

Insertions between the two hsp70 genes at 87A7 seem to be common in *D. melanogaster*. The 56H8 clone of Schedl *et al.*¹⁷ derives from this locus and contains a dispersed repetitive sequence between the two hsp70 genes¹⁸. Other changes in the structure of the 87A7 locus in chromosomes derived from natural populations seem to be due to independent events of this type (A.J.L.B., unpublished observations). After heat shock the inserted $\alpha\beta$ elements are transcribed from 87C1 but no other sites to give an abundant polyadenylated RNA^{7,16}. Thus the sequence rearrangements at 87C1 have resulted in activation of previously unexpressed $\alpha\beta$ sequences, although the transcripts do not seem to be essential for viability or fertility^{13,19}. This suggests that DNA rearrangements during evolution may be responsible for more rapid change in gene function than base substitution.

Nucleotide substitution at the hsp70 loci

We can study the evolution of the hsp70 loci using the restriction maps (Fig. 1) to estimate the rate of nucleotide substitution. This is possible because the three maps are independently derived and the two loci were distinguished using locus-specific probes,

649 site differences from the 87A and 87C loci of *Drosophila* species. The 5' and 3'-terminal *Sal*I sites are indicated and the distribution of restriction sites shown. Within a locus z elements are identical unless otherwise indicated. To estimate the probability of parallel change in a restriction site in both members of a pair of z elements (see text), we calculated the proportion of restriction sites conserved between homologous z elements in *D. melanogaster* and each of the other two species⁴². (Independent evidence suggests that *D. simulans* and *D. mauritiana* are more closely related to each other than either is to *D. melanogaster*¹⁴.) The four proportions were pooled to give an overall estimate of the mean number of restriction sites conserved of 0.828. The probability that a given site is different in two homologous z elements is thus $1 - 0.828 = 0.172$, and the probability that the same change occurred independently in a pair of sequences is $0.0297 (= 0.172^2)$. We have found six examples of parallel change in 36 comparisons of restriction site pairs ($P = 0.00062$). kb, Kilobase.



not restriction-site homology. Comparison of hsp70 genes from 87A and 87C in *D. melanogaster* has shown that the non-protein-coding regions have diverged much more than the protein-coding regions^{11,20}. Presumably, the latter are constrained by the need to retain protein function. We have therefore considered the nucleotide substitution of the region coding for hsp70 protein separately from that of the surrounding sequences (Table 1).

We can use the degree of nucleotide divergence to estimate the rate of substitution for the non-protein-coding regions. Carson²¹ has described a correlation in Hawaiian *Drosophila* between estimated divergence time and genetic distance based on frequencies of enzyme variants. The genetic distance between *D. simulans* and *D. melanogaster*²² would suggest a divergence time of about 1×10^6 generations. We estimate the degree of nucleotide substitution in non-coding DNA to be 0.03 per nucleotide (Table 1). This is equivalent to a rate of nucleotide divergence of the order of 3×10^{-8} per nucleotide per generation. Although estimations of divergence time are subject to large errors, we note that this rate is consistent with estimates for primates (approximately 1×10^{-8} per nucleotide per generation) obtained from the measurement of melting points of interspecific hybrid DNA molecules^{23,24}.

Intra-locus gene conversion between z elements

We have remarked previously that z elements from the same locus have similar restriction maps in *D. melanogaster* and suggested that this may imply an intra-locus correction mechanism⁸. We now have supporting data on z elements at six hsp70 loci. There are six examples of parallel change in a particular site in both z elements at the same locus [Fig. 3; *Pst*I (87A *D. melanogaster*); *Sal*I; *Bgl*I (87C *D. melanogaster*); *Pst*I (87C *D. simulans*), and the 3' *Xba*I sites at 87A and 87C in *D. melanogaster* which we discuss in more detail later]. We have estimated the probability of a restriction site changing independently in both members of a pair of z elements to be 0.0297 (see Fig. 3 legend). The probability of making such an observation six times or more in 36 comparisons is 0.00062. Clearly, the loss of restriction sites in both z elements at a locus cannot be independent and we believe that a mechanism is operating which corrects sequence differences that arise between members of a z pair. Three intra-locus restriction site differences occur at the 87C1 locus of *D. mauritiana*. Presumably these have arisen since the last conversion event.

Observations have been made on vertebrate globins which have led to similar conclusions. Duplicated human α -globin

sequences are considerably more similar than expected from the observed species divergence^{25,26}. These sequences are arranged in tandem, and it has been suggested that repeated rounds of unequal crossing-over could explain the observed sequence similarity. This is supported by the discovery of rare variants with three α -globin sequences in tandem²⁵. However, unequal crossing-over cannot act on inverted repeat sequences. We consider a different mechanism to explain the sequence homogeneity we have observed between z elements within hsp 70 loci.

Figure 4 illustrates our proposed mechanism, which is based on previous models, to explain gene conversion during meiosis^{27,28}. Occasional folding back of a chromosome would bring the two z elements in a locus close together in the same orientation. Strand exchange occurring between proximal and distal genes combined with base-mismatch repair could lead to a locus containing a pair of similar z elements. Although Fig. 4 shows symmetric strand displacement, asymmetric exchange will also lead to sequence transfer between proximal and distal genes. The extent of gene conversion will depend on the length of branch migration so that it need not lead to exact similarity between sequence pairs. Note that unequal crossing-over cannot explain such incomplete exchange. A possible consequence of this model is the inversion of the DNA region between the two genes if the two sets of strand breakages occur on opposite strands (Fig. 4C, D). A similar mechanism could also operate between tandem genes except that symmetric exchange would give rise to deletions. Indeed, an intra-chromosomal recombination mechanism in *Drosophila* has been proposed to explain the reversion of certain tandem duplications at the *Bar* and *Beadex* loci in the absence of meiotic recombination and sister-strand exchange²⁹⁻³¹. We note that gene conversion has recently been invoked to explain discontinuities in sequence conservation found between human γ - and α -globin genes³².

Presumably, the likelihood of gene conversion will depend on the distance between the two genes. Thus, one would expect the frequency of intra-locus exchange to be greatly diminished by the insertion of the large $\alpha\beta$ -containing region into the 87C1 locus of *D. melanogaster*. Sequence comparison between gene pairs at 87A7 would allow an estimate of the rate of exchange between two closely spaced hsp70 genes. If the insertion of the $\alpha\beta$ domain into the 87C1 locus causes a dramatic reduction of intra-locus conversion, the rate of nucleotide substitution between proximal and distal hsp70 genes will rise. Determination of the degree of sequence divergence between proximal and distal hsp70 genes at 87C1 in *D. melanogaster* will allow estimation of the time of insertion of the $\alpha\beta$ domain.

Sequence homogeneity between z elements from different loci

In a recent study³³, cloned hsp70 coding sequences derived from *D. melanogaster* were hybridized to the polytene chromosomes of the unrelated species *Drosophila pseudoobscura*. This revealed two linked sites of sequence homology at heat-induced chromosome puffs. These are located on chromosome 2 in *D. pseudoobscura* which corresponds to the chromosome arm (3R) in *D. melanogaster* that contains the hsp70 loci. This suggests that the duplication event which gave rise to two hsp70 loci is very ancient, the *obscura* and *melanogaster* lineages probably diverging in the early Oligocene some 35×10^6 yr ago³⁴. Despite this long separation the z elements of the two loci are very similar.

Comparison of the z elements in the three species studied shows no greater divergence between the two hsp70 loci than that between species at one locus (Fig. 3). In particular, only one restriction site differentiates z elements from the two loci of *D. simulans*. This suggests that a 'correction' mechanism is acting between the two hsp70 loci, despite their being about 500 kilobases apart. Further evidence of such sequence exchange comes from the distribution of *Xba*I sites in the z elements.

All z elements in *D. melanogaster* have an *Xba*I site some 300 bases upstream from the terminal *Sal*I site^{8,11,35}. This site is absent from all copies in *D. simulans* and *D. mauritiana*. Three alternative origins of this distribution can be envisaged: (1) the 3'

*Xba*I site was present in the 'ancestral' z element and was lost from all copies in the lineage which led to *D. simulans* and *D. mauritiana*. (2) The loss of this site is the result of site-specific DNA modification rather than a coincidental change in the primary structure of the DNA of all z elements in *D. simulans* and *D. mauritiana*. (3) The 3' *Xba*I site arose in a z element at one locus in *D. melanogaster*, spread through that locus by an intra-locus conversion mechanism and then to the other hsp70 locus by an inter-locus gene conversion event.

We have data which seem to rule out the first two possibilities. First, none of the slightly more distantly related species, *D. teisseri*, *D. yakuba* or *Drosophila erecta*, has a 3' *Xba*I site in its z elements (A.J.L.B., unpublished observations). According to the first scheme, one or more of these would also be likely to have retained this site, as they are no more closely related to *D. simulans* than to *D. melanogaster*¹⁴.

We have examined the possibility of base modification by analysing the z elements of *D. melanogaster*/*D. simulans* hybrid flies, and find that none of the sequences derived from *D. simulans* has a 3' *Xba*I site and all those from *D. melanogaster* are cut by *Xba*I (codominant inheritance). It is unlikely that the lack of a 3' *Xba*I site in species other than *D. melanogaster* is due to a DNA modification system, although we cannot completely exclude an enzyme active only on previously modified sites.

The third alternative remains: that gene conversion events occur between the two hsp70 loci as well as within a locus. These must be rarer than intra-locus exchanges because we have found only one hsp70 locus with dissimilar hsp70 genes whereas the z elements differ between 87A and 87C1 in all three species studied. The relative rarity of inter-locus exchange reflects the lower likelihood of synapsis between non-adjacent genes. Although we cannot directly determine the frequency of gene conversion between the two hsp70 loci, we can estimate the time of the most recent exchange. The non-protein-coding sequences in z at 87A7 and 87C1 have diverged by 15%^{10,20}. This greatly exceeds our estimate of the degree of nucleotide divergence in noncoding DNA between these species (Table 1), indicating that the last complete conversion between the sequences at 87A7 and 87C1 greatly predated the separation of *D. melanogaster* and *D. simulans*. The conversion of the 3' *Xba*I site in *D. melanogaster* that must have occurred after speciation illustrates that conversion events need only exchange parts of the hsp70 gene. Recently, Scherer and Davis have demonstrated³⁶ gene conversion in yeast between unlinked *his3* genes associated with the yeast transposon, *Ty1*. Evidence for gene conversion events between unlinked suppressor genes in *Saccharomyces pombe* has also been reported³⁷.

The best characterized precedent for gene conversion between distant loci is the switching of mating type in yeast where conversion of the mating-type locus *MAT* by normally silent copies at *HML* and *HMR* has been demonstrated genetically and biochemically^{38,39}. Attempts to detect extra-chromosomal mating-type sequences were unsuccessful³⁸. The alternative model involves conventional recombinational mechanisms and resembles an inter-locus version of our model for intra-locus exchange^{40,41}. We believe that the mechanisms of intra- and inter-locus conversion of hsp70 genes are similar and differ mainly in frequency. Our results clearly indicate that these frequencies are sufficient to be of considerable evolutionary significance.

We thank Drs W. Bender, D. Charlesworth, R. Holliday, C. Langley and J. Williams, and our colleagues in the laboratory for stimulating discussions and criticism of the manuscript. We also thank Drs M. Ashburner, J. P. W. Young and R. A. Voelker for fly stocks, Dr M. Meselson for plasmids pPW229.1 and pPW232.1, and Drs P. Schedl and A. Udvardy for plasmids 56H8 and cHS2S2.

Received 5 November 1980; accepted 23 January 1981.

1. Ashburner, M. & Bonner, J. J. *Cell* **17**, 241-254 (1979).
2. Livak, K., Freund, R., Schweber, M., Wensink, P. C. & Meselson, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5613-5617 (1978).

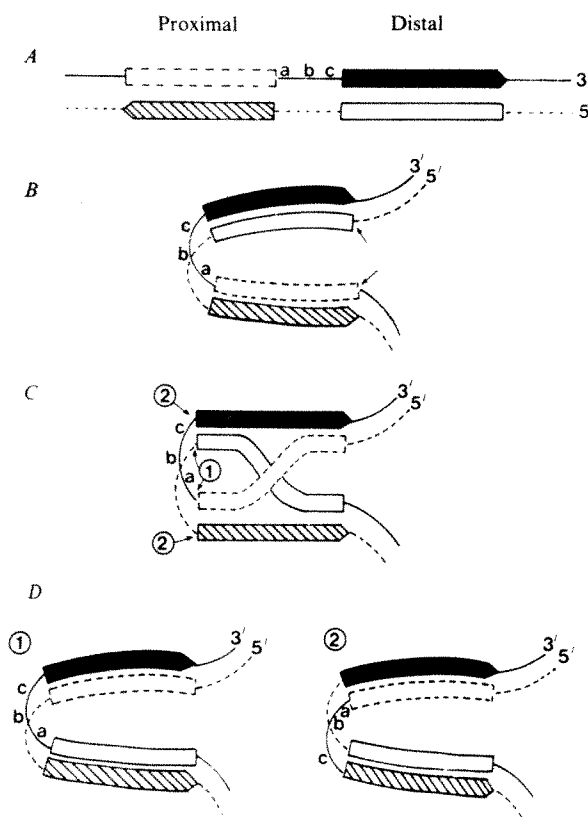


Fig. 4 Mechanism for intrachromosomal gene conversion. Filled boxes indicate coding sequences, transcribed in the direction of the arrow head, and empty ones the complementary sequences. Arrows indicate the sites of DNA strand breaks. The letters a, b and c indicate the orientation of the intergenic region. A, hsp70 gene alignment; B, double-stranded hairpin loop formation; C, strand displacement and exchange with two possible secondary sites of strand breakage; D, the two products. DNA repair will correct the base mismatch at the proximal and distal genes in the same direction in a proportion of these events. Note the inversion of the intervening markers in D(2). The mechanism depicted would operate as readily between sister chromatids as between sequences on the same chromatid. Indeed, a version which converts between paired homologues during meiosis can be envisaged.

3. Artavanis-Tsakonas, S., Schedl, P., Mirault, M.-E., Moran, L. & Lis, J. *Cell* **17**, 9–18 (1979).
4. Craig, E. A., McCarthy, B. J. & Wadsworth, S. C. *Cell* **16**, 575–588 (1979).
5. Ish-Horowicz, D. *et al.* *Cell* **17**, 565–571 (1979).
6. Lewis, M., Helmsing, P. J. & Ashburner, M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3604–3608 (1975).
7. Lis, J., Prestidge, L. & Hogness, D. S. *Cell* **14**, 901–919 (1978).
8. Ish-Horowicz, D., Pinchin, S. M., Schedl, P., Artavanis-Tsakonas, S. & Mirault, M.-E. *Cell* **18**, 1351–1358 (1979).
9. Mirault, M.-E., Goldschmidt-Clermont, M., Artavanis-Tsakonas, S. & Schedl, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5254–5258 (1979).
10. Holmgren, R., Livak, K., Morimoto, R., Freund, R. & Meselson, M. *Cell* **18**, 1359–1370 (1979).
11. Moran, L. *et al.* *Cell* **17**, 1–8 (1979).
12. Goldschmidt-Clermont, M. *Nucleic Acids Res.* **8**, 235–252 (1980).
13. Ish-Horowicz, D. & Pinchin, S. M. *J. molec. Biol.* **142**, 231–245 (1980).
14. Lemeunier, F. & Ashburner, M. *Proc. R. Soc. B* **193**, 275–294 (1976).
15. Southern, E. M. *J. molec. Biol.* **98**, 503–518 (1975).
16. Lis, J. T., Ish-Horowicz, D., Pinchin, S. M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
17. Schedl, P. *et al.* *Cell* **14**, 921–929 (1978).
18. Martin, G. *et al.* (in preparation).
19. Gausz, J. *et al.* *Genetics* **93**, 917–934 (1979).
20. Karch, F. & Torok, I. (in preparation).
21. Carson, H. L. *Nature* **259**, 395–396 (1976).
22. Ohnishi, S., Leigh Brown, A. J., Voelker, R. A. & Langley, C. H. *Genetics* (submitted).
23. Kohne, D. E. *Q. Rev. Biophys.* **3**, 327–375 (1970).
24. Benveniste, R. E. & Todaro, G. J. *Nature* **261**, 101–108 (1976).
25. Zimmer, E. A., Martin, S. L., Beverley, S. M., Kan, Y. W. & Wilson, A. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2158–2162 (1980).
26. Lauer, J., Shen, C.-K. J. & Maniatis, T. *Cell* **20**, 119–130 (1980).
27. Holliday, R. *Genet. Res.* **5**, 282–304 (1964).
28. Meselson, M. S. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 358–361 (1975).
29. Peterson, H. M. & Laughnan, J. R. *Proc. natn. Acad. Sci. U.S.A.* **50**, 126–133 (1963).
30. Laughnan, J. R. & Gabay, Molec. gen. Genet. **108**, 93–96 (1970).
31. Green, M. M. *Molec. gen. Genet.* **103**, 209–217 (1968).
32. Slightom, J. L., Blechl, A. E. & Smithies, O. *Cell* **21**, 627–638 (1980).
33. Pierce, D. A. & Lucchesi, J. C. *Chromosoma* **76**, 245–254 (1980).
34. Throckmorton, L. H. in *Handbook of Genetics* Vol. 3 (ed King, R. C.), 421–470 (Plenum, New York, 1974).
35. Torok, I. & Karch, F. *Nucleic Acids Res.* **8**, 3105–3123 (1980).
36. Scherer, S. & Davis, R. W. *Science* **209**, 1380–1384 (1980).
37. Hofer, F. *et al.* *Curr. Genet.* **1**, 45 (1979).
38. Hicks, J., Strathern, J. & Klar, A. J. S. *Nature* **282**, 478–483 (1979).
39. Nasmyth, K. A. & Tatchell, K. *Cell* **19**, 753–764 (1980).
40. Haber, J. E., Rogers, D. T. & McCusker, J. H. *Cell* **22**, 277–289 (1980).
41. Klar, A. J. S., McIndoo, J., Strathern, J. N. & Hicks, J. B. *Cell* **22**, 291–298 (1980).
42. Nei, M. & Li, W.-H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5269–5273 (1979).
43. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).

LETTERS

Detection of low order p-modes in brightness fluctuations of the Sun

Franz-Ludwig Deubner

Institut für Astronomie und Astrophysik, 8700 Würzburg, FRG

The discovery of radial and non-radial global oscillations of the Sun through various types of detector systems has opened the field of solar seismology. Of particular importance for studies of the deep interior of stars are the detection and analysis of low harmonic order pulsations such as the low order p-modes recently found on the Sun^{1,2} by line shift measurements. These pulsations have now also been detected in the solar brightness fluctuations. Through the reflected light of Uranus and Neptune the Sun was observed as though it were a distant star. The positive results suggest that broad-band photometry has the greatest potential for developing seismometry of main sequence stars.

High overtone pulsations of the Sun, both radial and non-radial (but with low degree l), discovered through line shift observations of the integrated solar disk by means of a potassium resonance spectrometer, have been reported by Claverie *et al.*¹ and confirmed with similar instrumentation operated at the South Pole station by Fossat² and his colleagues (also see ref. 3 for review). The excellent solar results available suggest that this type of measurement should be extended to main sequence stars and other 'nonvariables'. However, because of the lack of photons, the sensitivity of the kind of equipment necessary to select a narrow-band spectral region for line shift analysis is barely sufficient to resolve motions with anticipated amplitudes smaller than a few metres per second, even for the brightest stars⁴. The integrated power in the low order p-mode band has an amplitude of $\sim 2 \text{ m s}^{-1}$. However, decomposition into its individual modes requires a sensitivity better than 20 cm s^{-1} , as in the solar case. For this reason, and also because of the more common availability of good broad-band photometric devices, it is of interest to study the possibility of detecting comparatively short period fluctuations of the integrated intensity related to the p-mode oscillations.

Clearly, one does not hope to see the geometrical effect of a change of the radius by $\delta r/r \approx 1.5 \times 10^{-8}$. It is the enhancement by a factor of 1,000–2,000 of the temperature fluctuation compared with the fractional change of the radius^{3,5} which brings the expected brightness fluctuations into the range of $\delta I/I \geq 2 \times 10^{-5}$ for vertical photospheric motions with the

amplitude (20 cm s^{-1}) and period (5 min) of the p-modes. Compared with common scintillation amplitudes of the order of 10^{-3} mag the expected signals still seem to be much too small to be detected.

For a successful observation of non-radial stellar p-modes it is important to resolve and identify the individual modes, in addition to detecting the signal. To obtain a resolution in frequency sufficient to resolve the solar p-mode spectrum (that is, better than $30 \mu\text{Hz}$), observing sessions of at least 10 h, preferably without interruption, are required. The duration of this observation will then also help to suppress the noise level due to atmospheric scintillation.

A recent search for intensity fluctuations of solar origin in daylight observations of the integral sunlight⁶ was unsuccessful. I have previously obtained⁷ data amenable to similar analysis at the ESO station in Chile, to investigate the presence of fluctuations with periods longer than 10 min in the solar luminosity, as monitored by broad-band dual beam photometry of the brightness of Uranus and Neptune using a 50-cm telescope. Combining observations of 10 successive nights with a mean observing duration of 7.8 produced a r.m.s. noise level of 6×10^{-5} in these data. In the present investigation the spectral analysis was repeated applying a least squares sine wave fitting technique to the complete set of data, to avoid the awkward consequences of the temporal window in a straight Fourier transform with zeroes padded in the data gaps. Power estimates were computed at $(1 \text{ cycle}/14,400) \text{ min}^{-1}$ intervals. The power spectrum is still very noisy and gives only a vague impression of regularly spaced features. Fortunately, asymptotically the p-mode eigenfrequencies are spaced at regular intervals, and the size of this interval is well enough established⁸ to permit a 'superposed frequency' analysis analogous to the superposed epoch technique used to detect non-sinusoidal periodic fluctuations. A superposed frequency analysis applied to a velocity power spectrum, which is analogous to the procedure here, was recently published by Grec *et al.*²

The result of this procedure, averaging 15 frequency segments Δf of $135.42 \mu\text{Hz}$ each, covering the range 2.083–4.167 mHz, is shown in Fig. 1b, accompanied by the 'unfolded' spectral data on the same frequency scale in Fig. 1a. At one position the signal is well above the 3σ level of this mean power distribution, at two more positions the signal exceeds the 2σ level. All three features are accompanied by ghosts at 10 frequency intervals' distance due to the periodic structure of the window function. (The apparent asymmetry of the ghost pattern can be shown by numerical simulation to be due partly to the shape of the window transform and partly to the comparatively high noise level of the data.) The principal frequencies falling next to the mean period of 5 min are 3.301, 3.309 and 3.377 mHz.

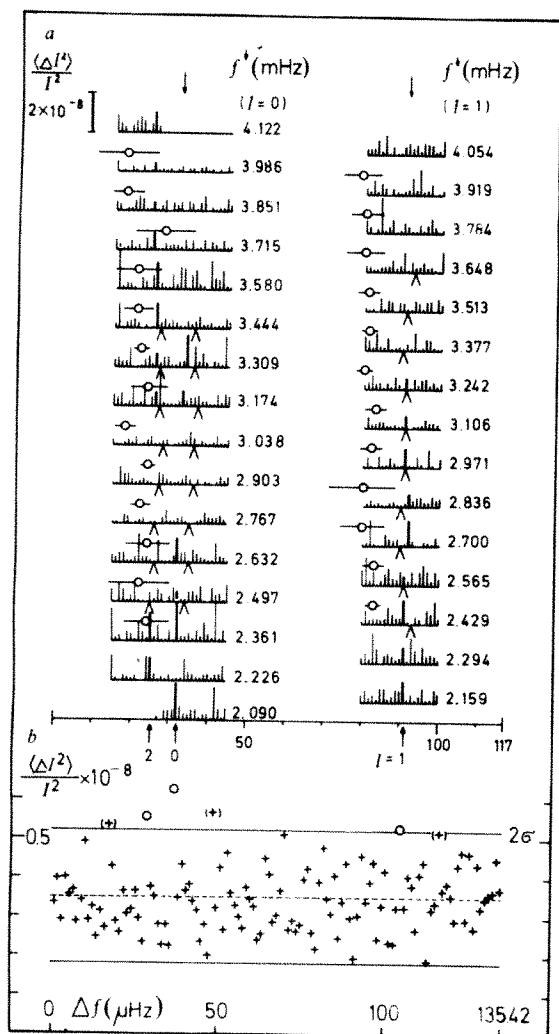


Fig. 1 Frequency spectrum of brightness fluctuations of the Sun. Spectral segments of length 135.42 μHz , shown in *a* are averaged to obtain the 'superposed' spectrum in *b*. Three peaks above the 2σ level are believed to represent sets of low order solar p-modes (ghost lines are bracketed). In *a*, the observed spectral frequencies which are believed to represent global p-modes are drawn with heavier lines. The frequency values given refer to the position within each segment marked by the arrows on top. The open dots with error bars in *a* are frequencies observed through line shift observations by Claverie *et al.*⁸. The arrow heads mark three sets of frequencies determined by Fossat (personal communication) from a 5 days continuous observation obtained at the South Pole. The numbered scale below the segmented spectrum indicates the 117 frequency intervals of $(1/14,400) \text{ min}^{-1}$ congruent with the folding interval of 135.42 μHz .

With the improved frequency resolution of this observation, one of the two sets of eigenfrequencies reported by Claverie *et al.*¹ is resolved into doublets at an interval of $\sim 8 \mu\text{Hz}$, similar to Fossat's² result of $10 \mu\text{Hz}$. Identification of these doublets with $l=0$ and $l=2$ modes is consistent with theoretical spectra published by Iben and Mahaffy⁹ and Christensen-Dalsgaard¹⁰. In detail, our observed frequencies agree quite satisfactorily with those obtained by Fossat (personal communication), but are systematically higher by 5–10 μHz than the values given by Claverie *et al.*⁸. To some extent, the accuracy of these latter data may suffer from insufficient spectral resolution. By standard statistical criteria, the significance of the three peaks emerging from the 2σ background of our superposed frequency analysis seems very marginal. We note, however, that of these peaks, two (with $l=1$ and 2) coincide within $\pm 2.25 \mu\text{Hz}$ with the most accurately observed superposed frequencies available. The probability of such a coincidence occurring by chance is 1.7×10^{-2} . Considering the statistical fluctuations of the amplitude estimates of a 10 days' data sample the mode identifications

given in Fig. 1 are consistent with the theoretical result that the mode numbers $l=0, 1$ and 2 are those most favoured by the spatial filter function both of full disk velocity and intensity observations¹¹. Finally, a phase difference of 3.5 yr in the solar activity cycle existing between measurements of intensity⁷ and velocity (E. Fossat, personal communication) should be considered when minor systematic frequency differences among these observations are confirmed.

After subtraction of the noise the observed mean power at the p-mode frequencies corresponds to a r.m.s. amplitude of $4\text{--}5 \times 10^{-5}$, in good agreement with our expectation. The improved signal-to-noise ratio obtained with stellar objects brighter than Uranus and Neptune, and with larger telescopes should facilitate the analysis by making the periodic structure stand out against the straight unfolded spectrum.

I thank G. Isaak and E. Fossat for unpublished results, and for exciting discussions.

Received 22 December 1980; accepted 12 February 1981.

1. Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. & Roca Cortes, T. *Nature* **282**, 591 (1979).
2. Grec, G., Fossat, E. & Pomerantz, M. *Nature* **288**, 541–544 (1980).
3. Gough, D. *Lecture Notes Phys.* **125**, 273 (1980).
4. Traub, W. A., Mariska, J. T. & Carleton, N. P. *Astrophys. J.* **223**, 583 (1978).
5. Hill, H. A., Rosenwald, R. D. & Caudell, T. P. *Astrophys. J.* **225**, 304 (1978).
6. Claverie, A., Isaak, G. R. & McLeod, C. P. *Proc. 14th ESLAB Symp.*, Scheveningen (in the press).
7. Deubner, F. -L. *Astr. Astrophys.* **57**, 317 (1977).
8. Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. & Roca Cortes, T. *Proc. 14th ESLAB Symp.* Scheveningen (1980).
9. Iben, I. Jr & Mahaffy, J. *Astrophys. J.* **209**, L39 (1976).
10. Christensen-Dalsgaard, J. & Gough, D. O. *Nature* **288**, 544 (1980).
11. Hill, H. A. in *The New Solar Physics* (ed. Eddy, J. A.) 135 (Westview Press, Colorado, 1978).

Solar neutrinos and turbulent diffusion

E. Schatzman

Observatoire de Nice, BP 252, Nice Cedex 06007, France

A. Maeder

Observatoire de Genève, 1290 Sauverny, Switzerland

Stellar rotation generates instabilities¹ which have not been taken into account in classical computations of stellar evolution. In particular, these instabilities may produce turbulent diffusion with important consequences over long time scales. Thus it has been shown^{2,3} that the loss of angular momentum which occurs in the external stellar layers due to stellar wind and magnetic braking⁴ leads to a redistribution of angular momentum. We show here that there are observational and theoretical reasons to suspect that a slow diffusion process due to a mild turbulence is operating in stellar interiors. Turbulent diffusion may have important consequences for the solar neutrinos problem and the (³He/⁴He) ratio at the solar surface.

The necessary and sufficient condition for the generation of turbulence has been often discussed. The presence of shear is a necessary but not sufficient condition; the Richardson criterion⁵ has also to be satisfied. But, when account^{6–8} is taken of the heat exchange between the moving fluid and the surrounding medium the Richardson criterion has to be replaced by the less stringent Townsend criterion or flux Richardson criterion. Eliassen and Kleinschmidt⁹ used the Townsend criterion to describe the maintenance of turbulence which is the same as that used by Zahn⁸. These conditions naturally differ from those of Layzer¹⁰, who has assumed that no heat exchange takes place during the motion of the fluid, a more stringent requirement.

The Rayleigh condition, states that, as long as $\delta\Omega/\delta z = 0$, the fluid is stable provided $\delta(\Omega\varpi^2)/\delta\varpi < 0$, Ω being the angular velocity and ϖ the distance to the axis. This last condition is violated if the star, like the Sun, is losing angular momentum. But as soon as $\delta\Omega/\delta z \neq 0$ the rotational motion is dynamically unstable¹¹. It has been shown that there is no meridional circulation which can be stable everywhere in a star^{12,13}.

where that the absence of turbulence in the radiatively stable region of the Sun would be remarkable¹⁴. We assume that instabilities associated with the shear flows in differentially rotating stars generate a mild turbulence, the effects of which may be treated¹⁵ as a diffusion process in stellar evolution.

Several astrophysical observations seem to support the mechanism of turbulent diffusion. Schatzman¹⁶ suggested that turbulence could inhibit gravitational (and radiative) sorting of the elements, and Baglin¹⁷ showed that meridional circulation for fast-rotating stars can induce a sufficiently high level of shear flow turbulence to avoid the sorting of elements. The Li deficiency in giants of spectral types F and G has been measured by Alschuler¹⁸. Assuming that these giants belong to the first ascending branch, and that their mass can be derived from Iben's¹⁹ evolutionary tracks, the Li deficiency can be explained by Li burning during the main sequence phase: Li is carried by turbulent diffusion from the outer layers of the star towards the region where it is nuclearly processed and destroyed. For masses $M \leq 1.6 M_{\odot}$, the empirical estimates of this rate of Li destruction on the main sequence²⁰ essentially agree with the theoretical values based on the turbulent diffusion model.

The ^{13}C excess in giants of the first ascending branch²¹, can be accounted for by turbulent diffusion²² outwards from the region where ^{12}C is nuclearly processed and partly transformed into ^{13}C .

Note that the equations for a slow diffusion process in stellar evolution are quite different from those for convective mixing. In particular, homogenization in finite volumes²³ does not properly mimic a slow diffusion process. This can be explained by a simple physical argument. The complete mixing over a fraction q_m of the solar mass²³⁻²⁶ generates a step function of the chemical composition, with a jump in the chemical composition at the mass fraction q_m . For solar models, with the constraint that the luminosity is $L = L_{\odot}$ for $t = 4,650$ Myr, we have $L \sim X_{1\text{en}}^2 T_c^4$, where $X_{1\text{en}}$ is some convenient average value of the hydrogen concentration X_1 . This means that $T_c \sim X_{1\text{en}}^{-1/2}$. With the power law approximation²⁷ $\phi_{\nu} \sim T_c^{20} X_4 (1 + X_1)^{-1}$, the reduction in the neutrino flux is given by

$$\frac{(\phi_{\nu})_{\text{mixed}}}{\phi_{\nu}} \approx \left(\frac{X_{1c}}{X_{1\text{en}}} \right)^{-10} \frac{(X_4)_{\text{mixed}}}{X_4} \frac{(1 + X_1)}{(1 + X_1)_{\text{mixed}}}$$

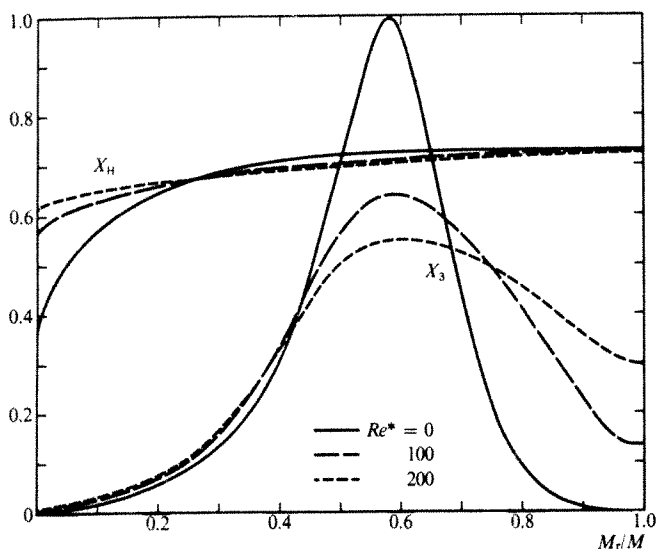


Fig. 1 Curves X_H give the hydrogen concentration as a function of the mass fraction. Turbulent diffusion brings hydrogen in the core of the Sun. Curves X_3 gives the ^3He concentration as a function of the mass fraction. Turbulent diffusion smoothes the ^3He peaks, with ^3He appearing at the surface and the concentration of ^3He increasing by an appreciable amount in the centre. The maximum X_3 corresponds to 2.8×10^{-3} . The increase of the concentration of the nuclear fuel in the centre is balanced by a decrease of the central temperature which produces a drastic reduction in the neutrino flux. Meanwhile, ^3He appears at the surface of the Sun.

The ratio $(X_{1c}/X_{1\text{en}})$ is always smaller for fully mixed models to $M = q_m M_{\odot}$ than for models where diffusion mixing has taken place. This is because the step function of the chemical composition is not a true description of the distribution of the chemical composition due to the diffusion process.

Schatzman¹⁵ and Genova and Schatzman²² demonstrated that a turbulent diffusion coefficient $D = Re^* \nu$, can be used where ν is the microscopic viscosity (molecular plus radiative) and Re^* is a critical Reynolds number of the order of 100. The changes of compositions in stellar models due to both the nuclear reactions and the turbulent transport may be expressed as:

$$\rho \frac{\partial X_1}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \rho D \frac{\partial X_1}{\partial r} + \rho A_1 m_H (-3R_{11} + 2R_{33} - R_{34})$$

$$\rho \frac{\partial X_3}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \rho D \frac{\partial X_3}{\partial r} + \rho A_3 m_H (R_{11} - 2R_{33} - R_{34})$$

$$\rho \frac{\partial X_4}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \rho D \frac{\partial X_4}{\partial r} + \rho A_4 m_H (R_{33} + R_{34})$$

X_1 , X_3 and X_4 are functions of r and t and represent the abundances, expressed in mass fraction, of ^1H , ^3He and ^4He respectively. ρ is the density and the R_{ij} terms are the nuclear reaction rates²⁸ for the species ij . An efficient matrix scheme for the diffusion equations²⁹ has been incorporated into our stellar evolution code.

Models with $Re^* = 200$, 100 and 0 (standard case) for an initial composition $X_1 = 0.73$, $X_4 = 0.25$ and $Z = 0.02$ have been computed. We note the following main results. As shown in Fig. 1, which gives the distribution of X_1 and X_3 in the Sun at $t = 4,600$ Myr, there is much more central hydrogen in the present Sun when diffusion is included. While the central X_1 is only 0.359 in the standard case, it is 0.569 with $Re^* = 100$. This excess of central X_1 evidently results¹⁵ from the inward transport due to a positive abundance gradient. Outside the central regions, the differences are smaller and of opposite sign. Overall, turbulent diffusion makes the X_1 distribution much flatter.

The internal distribution of ^3He shows a peaked distribution in the standard model (see Fig. 1). With turbulent diffusion, with $Re^* = 100$ and 200, the distribution of ^3He below $M_r/M = 0.4$ is still mainly determined by the rapid equilibrium between creation and destruction as in the standard case. Above $M_r/M = 0.4$ ($T < 8.3 \times 10^6$ K), the transport by turbulent diffusion tends to unify the ^3He distribution³⁰: the peak is reduced and ^3He is spread throughout the star. Figure 1 shows that the surface ratio $^3\text{He}/^4\text{He}$ is considerably increased by turbulent diffusion. An acceptable observational value of this ratio^{31,32} is $^3\text{He}/^4\text{He} \approx 4 \times 10^{-4}$. This value is $\sim 2-4$ times larger³³ than that of the protosolar material and implies a surface enrichment in ^3He during solar evolution. This result supports the occurrence of turbulent diffusion with a value of $Re^* \approx 60$.

The decrease of the mean molecular weight in the interior due to diffusion leads to a decrease of the central temperature and density and, in models with turbulent diffusion, the central concentration of ^3He is increased. Some characteristics of the models at $t = 4,600$ Myr are given in Table 1. We see that the neutrinos flux is smaller by a factor of 8.1 for $Re^* = 200$, 4.9 for $Re^* = 100$ and 2.6 for $Re^* = 50$ relative to the flux obtained in the standard case $Re^* = 0$.

The constraint on the solar luminosity usually fixes the helium content in the initial model; a change of X_4 by -0.01 reduces the luminosity by $\log L = -0.028$. Table 1 shows that diffusion models may fulfill the requirement on solar luminosity with a helium content differing only by 0.01–0.02 from that of the standard case.

Estimates have been made of the possibility of choosing a Z (heavy elements), X_4 (Helium) and Re^* to fulfill the three main requirements: the value of the solar luminosity, the ^3He concentration at the surface and the solar neutrino flux. As a whole, models with turbulent diffusion with Re^* in the range

Table 1 Properties of the standard model of the Sun and of the models with turbulent diffusion mixing at $t = 4,600$ Myr

Re^*	$\log T_c$	$(X_1)_c$	$(X_3)_c$	N_ν (SNU)	$(X_3/X_4)_s$	$\log L/L_\odot$
0	7.1944	0.3585	$8 \cdot E-6$	11.66	—	0.048
100	7.1681	0.5685	$2.1 \cdot E-5$	2.38	$1.57 \cdot E-3$	0.009
200	7.1601	0.6164	$2.6 \cdot E-5$	1.43	$3.38 \cdot E-3$	0.003

$X = 0.73$, $Y = 0.25$, $Z = 0.02$.

$(X_1)_c$ is the hydrogen concentration in the centre; $(X_3)_c$ is the ^3He concentration in the centre; $(X_3/X_4)_s$ is the surface concentration ratio of ^3He and ^4He at the surface.

50–100 seemed to fulfill the main solar constraints for a standard composition.

Our solution overcomes the difficulties of the low X_4 and low Z solar models³⁴. However, recent observations of the solar oscillations³⁵ have led Christensen-Dalsgaard and Gough³⁶ to describe solar models which agree with standard models. The difficulty³⁷ due to migration of heavy elements through a shallow convective zone can be avoided. The same model explains the depletion of lithium at the surface of the Sun.

From dimensional arguments alone we can see that the depletion of lithium would be of the order of $\exp[-(t_0/\tau_{\text{Li}})]$ and the depletion of Be would be of the order of $\exp[-(t_0/\tau_{\text{Be}})]$. τ_{Li} and τ_{Be} are of the order of $(\Delta r)_{\text{Li}}^2 D_{\text{turb}}^{-1}$ and $(\Delta r)_{\text{Be}}^2 D_{\text{turb}}^{-1}$ respectively, where $(\Delta r)_{\text{Li}}$ and $(\Delta r)_{\text{Be}}$ are the respective distances from the bottom of the hydrogen convective zone to the Li burning and Be burning regions. From the values of the Li depletion factor ($\approx 4.8 \times 10^2$) and of the Be depletion factor (≈ 2), and knowing the temperatures of the Li and Be burning regions ($T_{\text{Li}} \approx 2.4 \times 10^6$ K; $T_{\text{Be}} \approx 3.5 \times 10^6$ K), we can derive both where the hydrogen convective zone ends and the value of the turbulent diffusion coefficient (in fact, we do not obtain the bottom limit of the convective zone, but the point at which the penetrative convection T_{PCE} becomes inefficient):

$$\begin{aligned} T_{\text{PCE}} &= 2.19 \times 10^6 \text{ K} & R - r_{\text{PCE}} &= 2.06 \times 10^5 \text{ km} \\ T_{\text{Li}} &= 2.4 \times 10^6 \text{ K} & r_{\text{PCE}} - r_{\text{Li}} &= 0.2 \times 10^5 \text{ km} \\ T_{\text{Be}} &= 3.5 \times 10^6 \text{ K} & r_{\text{PCE}} - r_{\text{Be}} &= 1.24 \times 10^5 \text{ km} \end{aligned}$$

These values are compatible with those of Christensen-Dalsgaard and Gough³⁶. From the value of D_{turb} we obtain $Re^* \approx 85$ which is compatible with the values derived from the other solar constraints.

Meridional circulation³⁸ does not yield the same result. Meridional circulation drives the matter down from the bottom of the hydrogen convective zone to where the Li and Be are burning, and will not differentiate between Li and Be: the characteristic time of depletion being determined by the velocity of the circulation would be the same for both elements. Unless there was a very peculiar pattern of the meridional circulation, we can consider that this is a very strong argument in favour of turbulent diffusion and against meridional circulation²⁸.

We now comment on the role of turbulent diffusion in relation to the neutrino mass³⁹. Proposed values of the neutrino mass favour the suggestion^{40–42} that the neutrino oscillates between the Sun and the Earth. However, the non-zero mass of the neutrino does not necessarily imply an oscillation⁴³. It is therefore worth considering the consequence of change in the stellar models by introducing a simple physical effect and studying its implication for the Sun and the stars. It is difficult⁴⁴ to resolve the difference between predictions based on the solar models and observations by invoking neutrino oscillations only, if there are only three kinds of neutrinos coupled to each other. The probability of finding an electron neutrino after a time t is estimated by Bahcall and Frautschi⁴⁵ to be 0.5 and by Barger *et al.*⁴⁶ to be 0.62. With the latter value, this decreases for a standard model⁴⁴ the observable neutrino flux to 4.65 SNU, far above the observed flux.

How will the Sun evolve within the above framework? This question is particularly important in relation to stellar observations. Clearly, if turbulent diffusion were always maintained, a $1 M_\odot$ star would never become a red giant as the evolutionary time scale for such a star is larger than the diffusion time scale. The star would evolve towards the region of blue stragglers. To obtain an evolution towards the giant branch, turbulent diffusion has to be stopped. The progressively increasing μ -gradient during evolution probably⁸ has a stabilizing effect on the shear flow and consequently the turbulent diffusion would cease when the μ -gradient became large enough.

This suggests the existence of two groups of stars: (1) those with a very low differential rotation, which pursue a standard evolution; (2) those with a high degree of differential rotation, which evolve towards the blue-straggler region. There might also exist a wide range of intermediate cases for which the evolution towards the red-giant branch is only delayed and not stopped completely. This possibility seems very likely because cluster observations indicate⁴⁷ that the width of the main-sequence band is considerably larger than that given by standard models.

Although the evolution of the Sun is not very advanced, previous considerations suggest that this evolution would not necessarily follow route (1) above. In fact, we cannot yet estimate properly the period of rotation of the solar core. The minimum requirement for maintaining the turbulence is that the flux Richardson criterion is satisfied. This somehow gives the minimum angular velocity which might be present inside the Sun. Following Zahn⁸ this can be written

$$\varpi \frac{d\Omega}{d\varpi} \geq (\sigma Re_{\text{crit}}) \left(\frac{g}{H_p} \right) (\nabla_{\text{ad}} - \nabla_{\text{rad}})$$

where $\sigma = (\nu/\lambda_{\text{th}})$, ratio of viscosity to the thermal diffusivity is the Prandtl number, Re_{crit} the critical Reynolds number, H_p the pressure scale height and $\nabla = (d \log T / d \log P)$. This leads to a minimum value of Ω in the solar core, corresponding to a period of 2–3 days, in agreement with a vanishingly small oblateness. The turbulent diffusion coefficient for angular momentum therefore does not have a simple relationship with the turbulent diffusion coefficient for chemical species. Some feedback must exist, between loss of angular momentum, transport of angular momentum and turbulence which presumably stabilizes the critical Reynolds number Re^* around 100.

Received 3 November 1980; accepted 8 February 1981.

1. Tassoul, J. L. *Theory of Rotating Stars* (Princeton University Press, 1978).
2. Howard, L. N., Moore, D. W. & Spiegel, E. A. *Nature* **214**, 1297–1299 (1967).
3. Bretherton, F. P. & Spiegel, E. A. *Astrophys. J. Lett.* **153**, L77–8 (1978).
4. Schatzman, E. *Ann. Astrophys.* **25**, 18–29 (1962).
5. Richardson, L. F. *Proc. R. Soc. A* **97**, 354 (1920).
6. Townsend, A. A. *J. Fluid Mech.* **4**, 361 (1958).
7. Moore, D. W. & Spiegel, E. A. *Astrophys. J.* **139**, 48 (1964).
8. Zahn, J. P. *IAU Symp.* No 59, 185 (1974).
9. Eliassen, A. & Kleinschmidt, E. *Hdl d. Physik*, **48**, 45 (1957).
10. Layzer, D. in *Ionospheric Sporadic E* Vol. 2, 258 (Pergamon, Oxford, 1962).
11. Goldreich, P. & Schubert, G. *Astrophys. J.* **150**, 571 (1967).
12. Randers, G. *Astrophys. J.* **95**, 454 (1942).
13. Tassoul, J. L. *Theory of Rotating Stars* (Princeton University Press, 1978).
14. Kippenhahn, R. *IAU Symp.* No 66, 20 (1974).
15. Schatzman, E. *Astr. Astrophys.* **56**, 211–218 (1977).
16. Schatzman, E. *Astr. Astrophys.* **3**, 331–346 (1969).
17. Baglin, A. *Astr. Astrophys.* **19**, 45–50 (1972).
18. Alschuler, W. R. *Astrophys. J.* **195**, 649–660 (1975).
19. Iben, I. *Astrophys. J.* **142**, 1447 (1965); **147**, 650 (1967).
20. Scalo, J. M. & Miller, G. E. *Astrophys. J.* **239**, 953–960 (1980).
21. Lambert, D. L. *Mem. Soc. R. Sci. Liège* **9**, 405–415 (1976).
22. Genova, F. & Schatzman, E. *Astr. Astrophys.* **78**, 323–327 (1979).
23. Shaviv, G. & Salpeter, E. E. *Phys. Rev. Lett.* **21**, 1602–1605 (1968).
24. Ezer, D. & Cameron, A. G. W. *Astrophys. Lett.* **1**, 177 (1968).
25. Shaviv, G. & Beaudet, G. *Astrophys. Lett.* **2**, 17 (1968).
26. Bahcall, J. N., Bahcall, N. A. & Ulrich, R. K. *Astrophys. Lett.* **2**, 91 (1968).
27. Iben, I. Jr *Ann. Phys.* **54**, 164 (1969).
28. Fowler, W. A., Caughlan, G. R. & Zimmerman, B. A. *Rev. Astr. Astrophys.* **13**, 69 (1975).
29. Angrand, F. thesis, Univ. Pierre et Marie Curie, Paris (1980).
30. Schatzman, E. *Astrophys. Lett.* **3**, 139–140 (1969).
31. Geiss, J., Buhler, F., Cerutti, H., Eberhardt, P. & Filleux, Ch. *Apollo 16 Prel. Sci. Rep.* (NASA SP-315, 1972).
32. Hall, D. N. B. *Astrophys. J.* **197**, 509 (1975).
33. Geiss, J. & Bochsler, P. Preprint (1980).
34. Christensen-Dalsgaard, J., Gough, D. O. & Morgan, J. G. *Astr. Astrophys.* **73**, 121–128 (1979).

35. Grec, G., Fossat, E. & Pomerantz, M. *Nature* **288**, 541 (1980).
36. Christensen-Dalsgaard, J. & Gough, D. O. *Nature* **288**, 544 (1980).
37. Michaud, G. *Nature* **266**, 433 (1977).
38. Vauclair, S., Vauclair, G., Schatzman, E. & Michaud, G. *Astrophys. J.* **223**, 567–582 (1978).
39. Close, F. E. *Nature* **284**, 507–508 (1980).
40. Maki, Z., Nakagawa, M. & Sakato, S. *Prog. theor. Phys.* **28**, 870 (1962).
41. Pontecorvo, B. Z. *exp. theor. Phys.* **53**, 1717 (1967); *Soviet Phys. JETP* **26**, 984 (1968).
42. Gribov, V. & Pontecorvo, P. *Phys. Lett.* **28B**, 493–496 (1969).
43. Mohapatra, R. W. & Senjenovic, G. *Phys. Rev. Lett.* **44**, 912–916 (1980).
44. Bahcall, J. N. *et al.* *Phys. Rev. Lett.* **45**, 945 (1980).
45. Bahcall, J. N. & Frautschi, S. C. *Phys. Lett.* **29B**, 623 (1969).
46. Barger, V., Whisnaut, K. & Phillips, R. J. N. *Phys. Rev. D* **22**, 1636 (1980).
47. Maeder, A. *Astr. Astrophys.* **32**, 177 (1974).

Solvent inclusion and disorder effects on an isostructural series of organic conductors

G. J. Ashwell

Department of Chemistry, Sheffield City Polytechnic, Pond Street, Sheffield S1 1WB, UK.

Interest in quasi-one-dimensional organic conductors has been renewed by the discovery¹ of superconductivity in (tetramethyltetraselenafulvalene)₂PF₆, (TMTSF)₂PF₆. At ambient pressure (TMTSF)₂PF₆ undergoes a Peierls distortion at 18 K (ref. 2) whereas under modest pressure the properties remain metallic with the onset of superconductivity at 0.9 K (ref. 1). There is now an incentive to synthesize other organic metals in which the Peierls distortion is suppressed in the hope that these may also superconduct under pressure. So far (DEPE)(TCNQ)₄(H₂O)_x is the only example^{3,4} to show a monotonic increase in conductivity, characteristic of a metal, on cooling at ambient pressure to below 30 mK. This metallic behaviour has, however, been disputed⁵. I verify here the unique temperature dependence of conductivity of (DEPE)(TCNQ)₄(H₂O)_x and report the electrical and structural properties of the isostructural congeners (DMPP)(TCNQ)₄(H₂O)_x and (DRPA)(TCNQ)₄(H₂O)_x where R = M (methyl) and E (ethyl) (see Table 1 for full names).

Needle-shaped crystals, 2 × 0.1 × 0.01 mm, were prepared by a method reported previously⁴ for bipyridinium-TCNQ salts. The crystals were fibrous and frequently bowed indicating a weak coupling between chains. The stoichiometries were

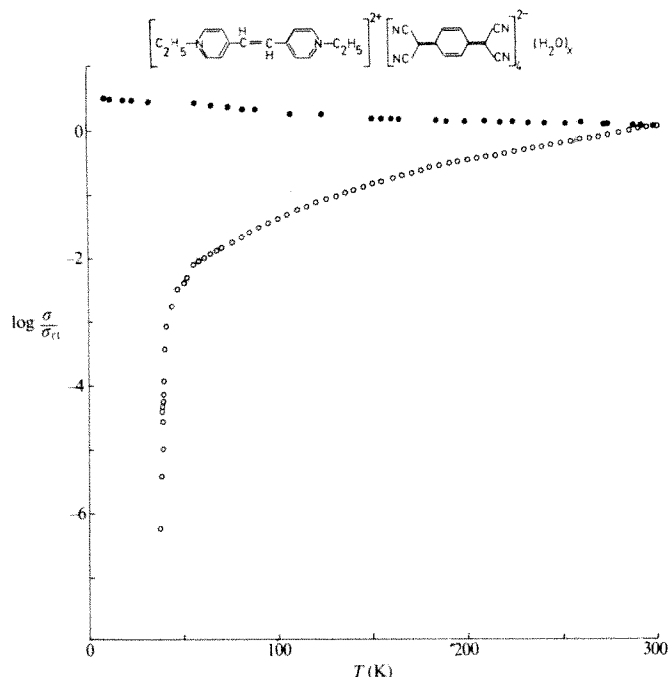


Fig. 1 Variation of the normalized conductivity with temperature of two crystals of (DEPE)(TCNQ)₄(H₂O)_x, one semiconducting (○) and the other metallic (●).

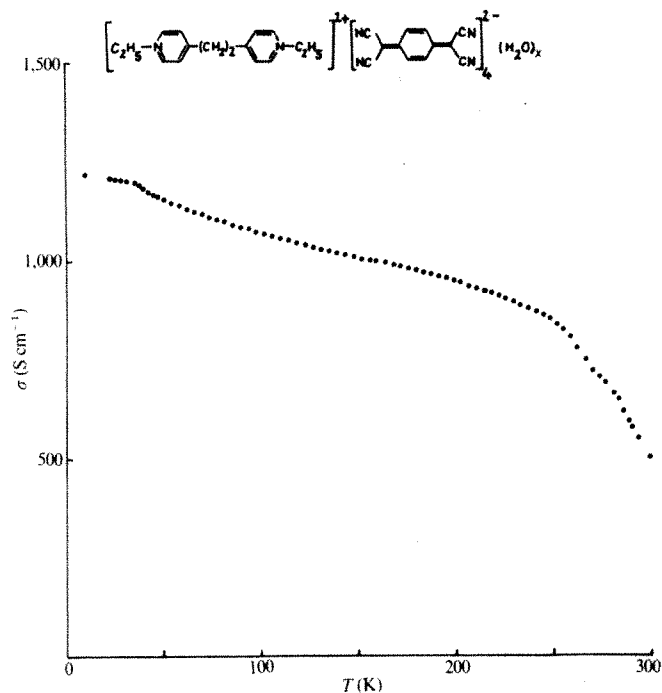


Fig. 2 Variation of the conductivity with temperature of (DEPA)(TCNQ)₄(H₂O)_x.

determined by elemental (C, H, N and O), spectral and thermogravimetric techniques (Table 1). The space group and unit cell parameters were obtained from oscillation and Weissenberg photographs using CuKα radiation (Table 2). Oscillation photographs about *b* showed strong subcell layer lines with *d*-spacings of 3.84–3.85 Å and weak layer lines with *d*-spacings of 17.4 Å in (DEPE)(TCNQ)₄(H₂O)_x, 17.1 Å in (DMPP)(TCNQ)₄(H₂O)_x and 7.68 Å in (DMPA)(TCNQ)₄(H₂O)_x. When the crystals were annealed in vacuum at ~370 K the weak layer lines could no longer be observed. In these conditions dehydration occurs with a contraction of the unit cell. The unit subcell parameters of (DMPP)(TCNQ)₄(H₂O)_x, refined on a Hilger and Watts four-circle diffractometer, were *a* = 12.93(2), *b* = 3.85(1), *c* = 27.50(2) Å, β = 108.1(2)°, *U* = 1,301 Å³ before annealing and *a* = 12.84(1), *b* = 3.84(1), *c* = 27.31(6) Å, β = 106.6(3)°, *U* = 1,290 Å³ after. In these quasi-one-dimensional conductors the weak layer lines are attributed to the H₂O lattice, commensurate in (DMPA)(TCNQ)₄(H₂O)_x and incommensurate in (DEPE)(TCNQ)₄(H₂O)_x and (DMPP)(TCNQ)₄(H₂O)_x. Variations in the intensities of the weak layer lines may indicate incomplete hydration and therefore a random occupancy of the available sites within a well defined H₂O lattice.

In (DEPE)(TCNQ)₄(H₂O)_x, (DMPP)(TCNQ)₄(H₂O)_x and (DMPA)(TCNQ)₄(H₂O)_x the cation lattice is inherently disordered which is characteristic of all highly conducting TCNQ salts with stoichiometries other than 1:1. The TCNQs stack plane-to-plane within homologous columns^{6,7} parallel to *b* with (1) a uniform mean interplanar spacing of 3.2 Å and (2) an exocyclic double bond to quinonoid ring overlap of adjacent molecules, with the longitudinal molecular staggering in a consistent direction throughout the stack. At ambient temperature the disorder inhibits a periodic coupling of the anion and cation chains and so suppresses the onset of a three-dimensional transition to a less conducting phase. At ~40 K the crystals undergo conductor-to-insulator transitions (Figs 1, 3) which are attributed to distortions of the quasi-one-dimensional TCNQ lattice. In these salts it is improbable that interactions between the conduction band electrons and the H₂O lattice, incommensurate in (DEPE)(TCNQ)₄(H₂O)_x and (DMPP)(TCNQ)₄(H₂O)_x, are sufficient to cause the Peierls distortions at

Table 1 Elemental, spectral and thermogravimetric data

Complex		C (%)	H (%)	N (%)	O (%)	$\epsilon_{395}/\epsilon_{842}$	Weight loss
(DEPE)(TCNQ) ₄ (H ₂ O) _x	Calculated*	71.65	3.52	23.51	1.31	2.2	1.5%
	Found	71.90	3.50	24.03	1.20	2.2	1.3%
(DMPP)(TCNQ) ₄ (H ₂ O) _x	Calculated†	71.31	3.57	23.77	1.36	2.2	
	Found	71.64	3.59	23.93	1.15	2.1	
(DMPA)(TCNQ) ₄ (H ₂ O) _x	Calculated‡	70.99	3.44	24.05	1.53	2.2	
	Found	71.01	3.23	24.80	1.36	2.1	

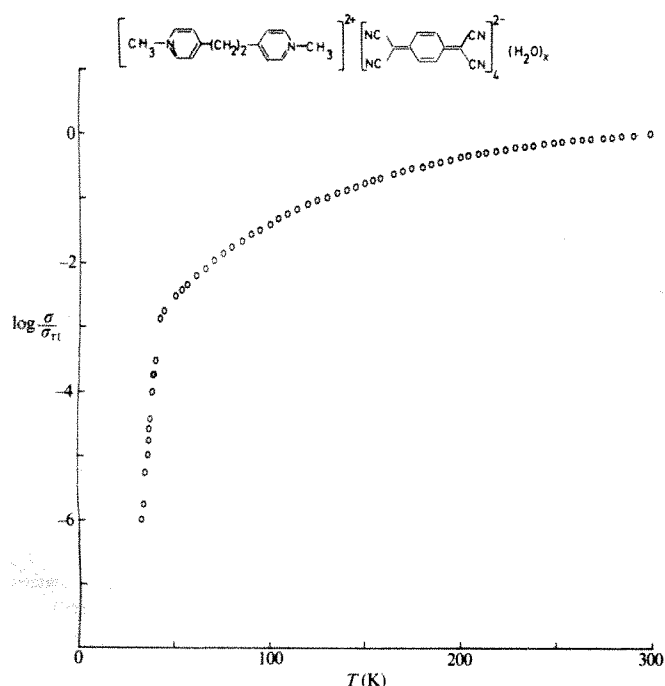
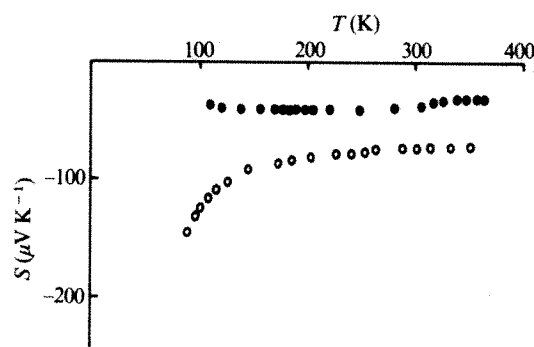
Values calculated by assuming: * $x = 0.88$; † $x = 0.90$ and ‡ $x = 1.0$. The non-stoichiometry is suggested by X-ray analysis; weak layer lines with d -spacings of $b/0.22$ in (DEPE)(TCNQ)₄(H₂O)_x and $b/0.225$ in (DMPP)(TCNQ)₄(H₂O)_x are attributed to an incommensurate H₂O lattice.

(DEPE)(TCNQ)₄(H₂O)_x = [1,2-bis(1-ethyl-4-pyridinium)ethylene]²⁺(7,7,8,8-tetracyanoquinodimethanide)₄²⁻(H₂O)_x. (DMPP)(TCNQ)₄(H₂O)_x = [1,3-bis(1-methyl-4-pyridinium)propane]²⁺(7,7,8,8-tetracyanoquinodimethanide)₄²⁻(H₂O)_x. (DMPA)(TCNQ)₄(H₂O)_x = [1,2-bis(1-methyl-4-pyridinium)ethane]²⁺(7,7,8,8-tetracyanoquinodimethanide)₄²⁻(H₂O)_x.

~40 K. If they were we should expect the transition in (DMPA)(TCNQ)₄(H₂O)_x to occur at a higher temperature, as the H₂O lattice is commensurate in this salt. Instead I suggest that the distortions arise, in part, from an ordering of the cation lattice on cooling.

The conductor-to-insulator transitions at 40 K have been suppressed, in a few instances only, by annealing the hydrated crystals in vacuum (10⁻⁴ torr). The conductivities increased on dehydration. For (DEPE)(TCNQ)₄(H₂O)_x, shown in Fig. 1, the conductivity decayed on thermal recycling but, nevertheless, remained metallic. For (DEPA)(TCNQ)₄(H₂O)_x, shown in Fig. 2, the conductivity was reproducible for two cooling cycles (300–10 K) and increased monotonically on cooling to below 10 K. In this low temperature region the properties depend on the degree of disorder which, in the limit, can obscure the Peierls transition by, for example, introducing a substantial density of states into the band gap as exemplified by recent work on (dimethyldiphenylphosphonium)⁺(TCNQ)₂⁻ (ref. 8). The experimental data reflect an increase in the disorder on dehydration and a decrease on thermal recycling.

The electrical properties of (DEPE)(TCNQ)₄(H₂O)_x, (DMPP)(TCNQ)₄(H₂O)_x, (DMPA)(TCNQ)₄(H₂O)_x and (DEPA)(TCNQ)₄(H₂O)_x depend on the outgassing procedure.

**Fig. 3** Variation of the normalized conductivity with temperature of (DMPA)(TCNQ)₄(H₂O)_x.**Fig. 4** Variation of the thermoelectric power with temperature of two compacted pellets of (DEPE)(TCNQ)₄(H₂O)_x, one hydrated (○) and the other dehydrated (●).

Crystals annealed in vacuum at 300–400 K have variable properties which may be attributed to the onset of decomposition, increased disorder and to loss of solvent from the lattice. Five crystals have shown metallic behaviour with $\sigma_{300K} = 100$ –500 S cm⁻¹ along b . Most were semiconducting with $\sigma_{300K} = 50$ S cm⁻¹. The metallic crystals are believed to be more disordered and less hydrated than the semiconducting crystals. For a correlated semiconductor the thermoelectric power, S , is given⁹ by

$$S = -k/e \ln[2(1-\rho)/\rho] \mu V K^{-1}$$

where ρ is the ratio of carriers to available TCNQ sites. Assuming that charge transfer is commensurate with the chemical formulae presented for the isostructural series, $\rho = 0.5$ and $S = -59.8 \mu V K^{-1}$. For (DEPE)(TCNQ)₄(H₂O)_x this theoretical value is consistent with the experimental data at saturation for the hydrated sample only (Fig. 4). The semiconducting behaviour probably arises from a pinning of the charge density waves by the H₂O 'impurities'. When the pinning is destroyed, by dehydration, the conductivity is enhanced and may become metallic. Thus in the dehydrated sample the thermoelectric

Table 2 Unit subcell parameters of the strong layer lines and d -spacings of the weak layer lines of the isostructural series of conducting TCNQ salts

Complex	Space group	a (Å)	b (Å)	c (Å)	β (°)	d -spacing (Å)
(DEPE)(TCNQ) ₄ (H ₂ O) _x	P2 ₁ /c	12.72	3.84	27.67	107.8	17.4
(DMPP)(TCNQ) ₄ (H ₂ O) _x	P2 ₁ /c	12.93	3.85	27.50	108.1	17.1
(DMPA)(TCNQ) ₄ (H ₂ O) _x	P2 ₁ /c	12.90	3.84	27.34	107.4	7.68

The weak layer lines are attributed to the H₂O lattice.

e.m.f. is significantly lower and the properties are no longer consistent with those of a correlated semiconductor. On dehydration there is no significant change in the TCNQ stacking characteristics.

All four salts have crystallographically ordered anhydrous modifications. In these the TCNQs are stacked plane-to-plane in stoichiometric groups of four, in a direction parallel to the length of the cation, with no direct overlap between adjacent tetrads. This periodic distortion arises from a periodic coupling of the anion and cation chains. The conductivities are significantly lower with $\sigma_{300K} \sim 0.002 \text{ S cm}^{-1}$ along b , the direction of stacking. Conversely in the crystallographically disordered modifications the periodic interchain coupling is destroyed by cation disorder and homologous columns of TCNQs occur. In both types of modification the electrical properties are consistent with those of related complexes in which the TCNQ stacking characteristics are similar.

I thank Dr S. C. Wallwork for the refined subcell parameters of $(\text{DMPP})(\text{TCNQ})_4(\text{H}_2\text{O})_x$ and Dr G. Owen for helpful discussions. I also thank G. H. Cross for assistance with the conductivity measurements and the Chemical Society Research Fund and the SRC for support.

Received 14 November 1980; accepted 9 February 1981.

1. Jerome, D., Mazaud, A., Ribault, M. & Bechgaard, K. *J. phys. Lett., Paris* **41**, L95 (1980).
2. Bechgaard, K., Jacobsen, C. S., Mortensen, K., Pedersen, H. J. & Thorup, N. *Solid St. Commun.* **33**, 1119 (1980).
3. Ashwell, G. J., Eley, D. D. & Willis, M. R. *Nature* **259**, 201 (1976).
4. Ashwell, G. J. *Phys. St. Solidi* **86b**, 705 (1978).
5. Somoano, R. B. *et al. Phys. St. Solidi* **81b**, 281 (1977).
6. Ashwell, G. J. *Proc. int. Conf. on Low Dimensional Synthetic Metals*, Helsingor (1980).
7. Wallwork, S. C. *Proc. Conf. on Organic and Biological Semiconductors*, Nottingham (1980).
8. Ashwell, G. J., Allen, D. W., Cross, G. H., Kennedy, D. A. & Nowell, I. W. (in preparation).
9. Chaikin, P. M. & Beni, G. *Phys. Rev.* **B13**, 647 (1976).

High abundances of long-chain normal alkanolic acids in Antarctic soil

Genki Matsumoto*, Tetsuya Torii†
& Takahisa Hanya*

* Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, Fukazawa 2-1-1, Setagaya-ku, Tokyo 158, Japan
† Chiba Institute of Technology, Tsudanuma, Narashino, Chiba Prefecture 275, Japan

Studies of geochemical samples have shown that fatty acids are widely distributed in sediments ranging in geological age from Recent to Precambrian¹⁻⁹. Most *n*-alkanoic acids found in soils and Recent sediments show similarities in their distributions, such as a marked predominance of even-carbon numbered acids and a bimodal distribution with two maxima in carbon chain length of C_{14} to C_{18} and C_{24} to C_{28} (refs 1, 2, 4-7). Long-chain *n*-alkanoic acids ($\geq \text{C}_{20}$) predominate in the waxes of vascular plants^{10,11}, whereas short-chain *n*-alkanoic acids ($< \text{C}_{20}$) occur in the lipids of all living organisms. Long-chain *n*-alkanoic acids found in soils and Recent sediments are thus believed to arise from vascular plants¹²⁻¹⁴. In Antarctica, vascular plants are only found in the Antarctic Peninsula¹⁵. Organic constituents on this continent are thus expected to be considerably different from those in the temperate and tropical zones. We report here the presence of very long-chain *n*-alkanoic acids extending to C_{40} , having a small even-carbon predominance, and comprising a large proportion of the fatty acids in Antarctic soils. They apparently originate in the Beacon Sandstone.

During the 1967-77 austral summer, soil samples were collected from the dry valleys of Victoria Land in Antarctica (Fig. 1) with a small stainless-steel scoop, wrapped in Teflon sheet and stored at temperatures below 0°C until analysis. The

procedures for fatty acid analysis have been described in detail elsewhere¹⁶. Briefly, wet samples ($\sim 50 \text{ g}$) were extracted with ethyl acetate after saponification with 0.5 M potassium hydroxide/methanol solution. Each extract was chromatographed on a silica gel column ($160 \times 4\text{-mm}$ i.d., 100 mesh, 5% water). After elution with two column volumes of hexane, the eluate (fatty acid fraction) obtained with three column volumes of benzene/ethyl acetate (95/5) was methylated with 15% boron trifluoride/methanol solution (3 ml, 80°C , 2 h). The structures of the methyl esters of fatty acids were determined using a combined Shimadzu LKB 9000 gas chromatograph-mass spectrometer in conditions given in Fig. 2.

Fatty acids were identified by comparing their retention times and mass spectra with those of authentic compounds and homologues. Quantitative acid data were obtained by measuring the peak area on the mass fragmentogram or gas chromatogram (total ion current). A blank test to check contamination throughout the experiments indicated possible contaminants (Fig. 2), but they were present in such low amounts that they did not affect the analytical results for long-chain *n*-alkanoic acids. Based on four replicate addition experiments using authentic compounds, the percentage recoveries of *n*-alkanoic acids, C_{14} , C_{16} , C_{18} , C_{20} and C_{22} , were 88 (s.d. 4.0), 91 (8.0), 87 (3.5), 88 (3.7) and 107 (13), respectively.

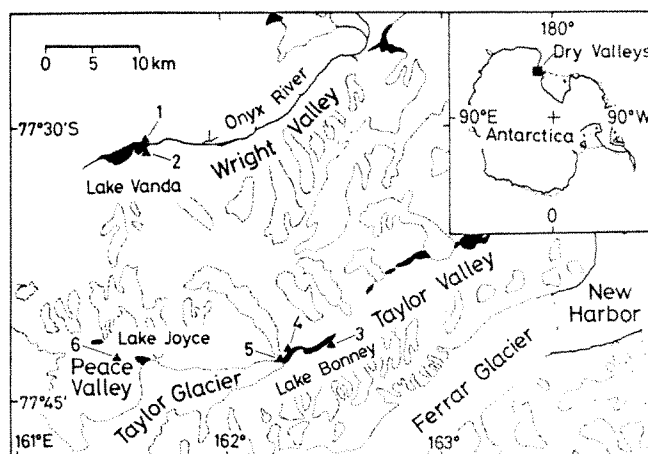


Fig. 1 Sampling locations in the dry valleys of Victoria Land in Antarctica. 1, Vanda-1; 2, Vanda-2; 3, Bonney, east lobe; 4, Bonney, west lobe-1; 5, Bonney, west lobe-2; 6, Joyce.

The analytical results from six soil samples are summarized in Table 1. The total acid content of the soils is fairly low, ranging from 0.16 to $3.7 \mu\text{g per g}$ dry soil, presumably reflecting the low organic carbon contents ($0.14\text{--}0.71 \text{ mg C per g}$)¹⁶. *n*-Alkanoic acids ranging from C_8 to C_{40} were found with a bimodal distribution maximizing at C_{16} and C_{28} , together with branched (iso and ante-iso) and unsaturated acids (Fig. 2). As far as we know, this is the first report of the occurrence of such very long-chain *n*-alkanoic acids ($> \text{C}_{38}$) in soils and modern and ancient sediments.

Surprisingly, long-chain *n*-alkanoic acids comprised most of the fatty acids, ranging from 30.2 to 81.9% with an average of $53.8 \pm 18.7\%$ at 90% confidence limits, with a small even-carbon predominance. Thus the ratios of long- to short-chain *n*-alkanoic acids are very high (2.7 ± 1.7) and comparable with the value estimated for a forest soil (2.3)¹¹. The carbon preference indices, CPI_{AL} , CPI_{AH} and CPI_{AT} , calculated over the ranges $\text{C}_{13}\text{--}\text{C}_{19}$, $\text{C}_{19}\text{--}\text{C}_{33}$ and $\text{C}_{13}\text{--}\text{C}_{33}$ according to Kvenvolden¹⁴, ranged from 4.5 to 13, 1.4 to 3.6 and 1.7 to 5.7, respectively. In general, these values are lower than those in common biological materials. In particular, the CPI_{AH} and CPI_{AT} values of the three Bonney samples are extremely low, and are thought to be the smallest values ever reported for soils, although CPI_{A} values close to unity have been found in some

Table 1 Fatty acids found in soils collected from the dry valleys of Victoria Land in Antarctica

Locality	Nature of soil	Range	Maximum at carbon number*	Concentration (μg per g)	Composition (%)				Ratio long/short	CPI _{AL}	CPI _{AH}	CPI _{AT}
					Short	Long	Branched	Un-saturated				
Wright Valley												
Vanda-1	OY, silt	C ₁₀ -C ₃₂	C ₁₆ , C ₂₄	3.7	31.1	31.6	5.4	31.9	1.0	13	3.6	5.7
Vanda-2	DYB, sand	C ₈ -C ₃₄	C ₁₆ , C ₂₈	2.4	30.4	46.7	9.4	13.5	1.5	7.8	3.2	4.3
Taylor Valley												
Bonney, east lobe	GO, sand	C ₈ -C ₄₀	C ₁₆ , C ₂₈	0.94	17.1	52.2	5.8	24.9	3.1	6.1	1.7	2.2
Bonney, west lobe-1	OY, sand	C ₈ -C ₄₀	C ₁₆ , C ₂₈	0.43	16.7	80.2	1.3	1.8	4.8	4.8	1.4	1.7
Bonney, west lobe-2	OY, sand	C ₈ -C ₃₇	C ₁₆ , C ₂₈	0.91	15.5	81.9	1.4	1.2	5.3	4.5	1.5	1.8
Pearce Valley												
Joyce	YB, sand	C ₈ -C ₃₅	C ₁₆ , C ₂₈	0.16	56.1	30.2	9.0	4.7	0.54	5.3	1.5	3.1
Average, 90% confidence limits.				1.4±1.1	27.8±12.8	53.8±18.7	5.4±2.9	13.0±10.6	2.7±1.7	6.9±2.6	2.2±0.81	3.1±1.3

Carbon preference indices, CPI_{AL}, CPI_{AH} and CPI_{AT}, values were calculated from the carbon chain lengths of C₁₃-C₁₉, C₁₉-C₃₃ and C₁₃-C₃₃, respectively. Short, short-chain *n*-alkanoic acids (C₁₃-C₁₉). Long, long-chain *n*-alkanoic acids ($\geq$ C₂₀). Branched, branched acids (iso and ante-iso, C₁₃-C₁₇). Unsaturated, unsaturated acids (C₁₆ and C₁₈). OY, olive yellow; DYB, dull yellowish brown; sand, coarse and fine sands; GO, greyish olive. YB, yellowish brown.

* The major distribution maximum is underlined.

sedimentary rocks and shales^{1,2,4}. However, we found no long-chain *n*-alkanoic acids with small CPI_A values, which make up a high proportion of the fatty acids in lake waters, sediments and epibenthic algae, that is, blue-green algae (mostly *Oscillatoria* spp., *Phormidium* spp. and *Calothrix* spp.), in samples collected from Lakes Vanda, Bonney and Joyce^{16,17}.

Little is known about the occurrence of long-chain *n*-alkanoic acids other than in the waxes of vascular plants^{10,11} and bees waxes^{18,19}. There are no vascular plants and bees in the areas studied and these long-chain *n*-alkanoic acids are not due to contamination during analysis. There seem to be at least three possible sources of these unusual *n*-alkanoic acids.

(1) An input of aeolian dust, because air-borne pollutants such as dichlorodiphenyl-trichloroethane are found in Antarctic snow²⁰. Some aeolian dusts collected at sea near land areas contain long-chain *n*-alkanoic acids in large relative abundances, but their CPI_A values (6.1-22.0) are considerably higher than ours²¹. Furthermore, the fatty acid ($\geq$ C₁₀)/*n*-alkane ($\geq$ C₁₅) ratios of the aeolian dusts (0.088-0.50)²¹ are much lower than those of the Antarctic soils (31-110, ref. 16; Matsumoto *et al.*, unpublished results). Thus aeolian dust is not the main source of our fatty acids.

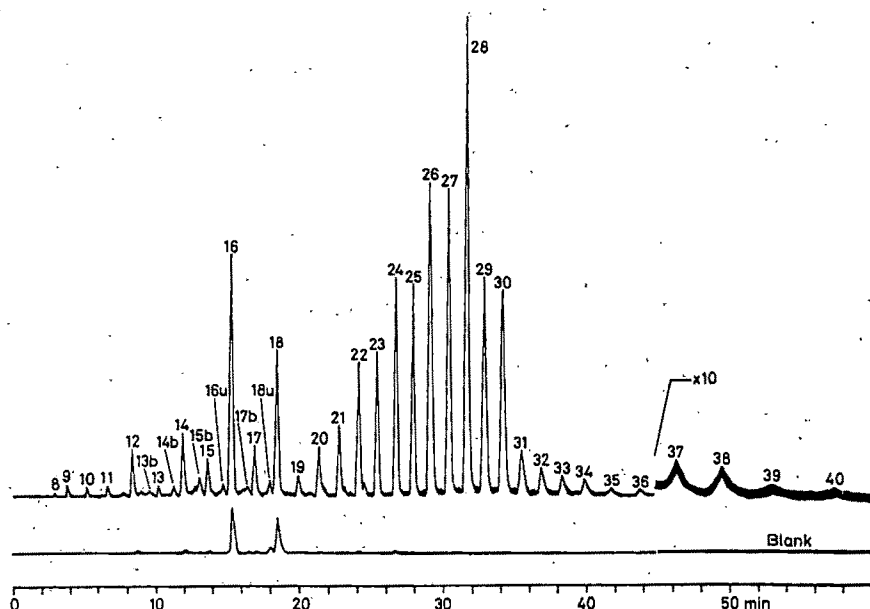
(2) Microbial oxidation of *n*-alkanes, because many kinds of microorganism including bacteria and fungi are found in soils in the dry valley areas²². Although *n*-alkanes ranging from C₁₅ to C₃₇, with the most dominant peak at C₂₃ or C₂₇, were found at very low total amounts, between 0.0045 and 0.055 $\mu\text{g per g}$ dry

soil (ref. 16; Matsumoto *et al.*, unpublished results), they could not give our fatty acid profile after microbial oxidation (Fig. 2, Table 2). Thus *n*-alkanes are also an unlikely source.

(3) Glacial erosion of sedimentary rocks containing organic materials may have contributed long-chain *n*-alkanoic acids or precursors, as long-chain *n*-alkanoic acids having small CPI_A values, and high long/short chain ratios have been found in sedimentary rocks and shales of Palaeozoic to Cenozoic age from the US^{1,2,4}. In our case, the dry valley areas are mainly composed of sedimentary, metamorphic and igneous rocks of Precambrian to Cretaceous age. It is well known that the Beacon Sandstone, composed of sandstone, siltstone, conglomerate and carbonaceous beds, and including various fossils such as coals and silicified woods, is widely distributed in the dry valley areas²³. Much of the sedimentary organic carbon in the Ross Sea is reported to be derived from the rocks being eroded by glaciers on the Antarctic continent²⁴. It is therefore likely that the long-chain *n*-alkanoic acids found in Antarctic soils come from the Beacon Sandstone. On the other hand, the short-chain acids, which are similar to those of common biological materials, can be mainly attributed to recent biological sources. Further samples of Antarctic sedimentary rocks including fossils will be collected for the analysis of organic constituents.

We thank Mr Y. Tanaka, Chiba Institute of Technology, the Antarctic Division, DSIR, New Zealand, the US NSF and the National Institute of Polar Research, Japan, for their support in collecting samples.

Fig. 2 Mass fragmentogram (m/e 74) of the fatty acid fraction obtained from soil sample collected from Victoria Land in Antarctica (Bonney, west lobe-1). Conditions: silanized glass column (200 cm $\times$ 3 mm i.d.) packed with 1% silicone OV-1 on 80-100-mesh Chromosorb W AW DMCS; temperature, injector 300 °C, column programmed from 100 to 290 °C at 6 °C min⁻¹, molecular separator 300 °C, ion source 330 °C; carrier gas, helium at 30 ml min⁻¹; acceleration voltage, 3.5 kV; electron energy, 20 eV; multiplier gain, 3. Arabic figures at the peaks indicate the carbon chain length of fatty acids. 13b, 14b, 15b and 17b are branched (iso + ante-iso) C₁₃, C₁₄, C₁₅ and C₁₇ acids, respectively. 16u and 18u are unsaturated C₁₆ and C₁₈ acids, respectively.



Received 25 March 1980; accepted 10 February 1981.

1. Kvenvolden, K. A. *Nature* **209**, 573–577 (1966).
2. Kvenvolden, K. A. *J. Am. Oil Chem. Soc.* **44**, 628–636 (1967).
3. Han, J. & Calvin, M. *Nature* **224**, 576–577 (1969).
4. Kvenvolden, K. A. in *Advances in Organic Geochemistry 1966* (eds Hobson, G. D. & Speers, G. C.) 335–366 (Pergamon, Oxford, 1970).
5. Brown, F. S., Baedecker, M. J., Nissenbaum, A. & Kaplan, I. R. *Geochim. cosmochim. Acta* **36**, 1185–1203 (1972).
6. Eglinton, G., Maxwell, J. R. & Philp, R. P. in *Advances in Organic Geochemistry 1973* (eds Tissot, B. & Bienner, F.) 941–961 (Éditions Technip, Paris, 1974).
7. Matsuda, H. & Koyama, T. *Geochim. cosmochim. Acta* **41**, 777–783 (1977).
8. Van Vleet, E. S. & Quinn, J. G. *Geochim. cosmochim. Acta* **43**, 289–303 (1979).
9. Volkman, J. K., Johns, R. B., Gillan, F. T. & Perry, G. J. *Geochim. cosmochim. Acta* **44**, 1133–1143 (1980).
10. Kolattukudy, P. E. *Lipids* **5**, 259–275 (1970).
11. Matsuda, H. & Koyama, T. *Geochim. cosmochim. Acta* **41**, 1825–1834 (1977).
12. Morrison, R. I. in *Organic Geochemistry. Methods and Results* (eds Eglinton, G. & Murphy, M. T. J.) 558–575 (Springer, Berlin, 1969).
13. Cranwell, P. A. *Chem. Geol.* **14**, 1–14 (1974).
14. Brooks, P. W. *et al. Chem. Geol.* **18**, 21–38 (1976).
15. Greene, S. W. *et al. in Terrestrial Life of Antarctica*, Antarctic Map Folio Series 5 (ed. Bushnell, V. C.) 1–24 (American Geographic Society, New York, 1967).
16. Matsumoto, G., Torii, T. & Hanya, T. *Mem. natl. Inst. Polar Res. Spec. Issue No. 13*, 103–120 (1979).
17. Matsumoto, G. & Hanya, T. *Antarctic Rec.* **58**, 81–88 (1977).
18. Holloway, P. J. *J. Am. Oil Chem. Soc.* **46**, 189–190 (1969).
19. Tulloch, A. P. & Hoffman, L. L. *J. Am. Oil Chem. Soc.* **49**, 696–699 (1972).
20. Peel, D. A. *Nature* **254**, 324–325 (1975).
21. Simoneit, B. R. T. & Eglinton, G. in *Advances in Organic Geochemistry 1975* (eds Campos, R. & Goñi, J.) 415–430 (Enadinsa, Madrid, 1977).
22. Cameron, R. E., King, J. & David, C. N. in *Antarctic Ecology Vol. 2* (ed. Holdgate, M. W.) 702–716 (Academic, London, 1970).
23. Warren, G. & Gunn, B. M. in *Geology of Mawson-Mulock Area, Victoria Land, Antarctica* (Dept of Scientific and Industrial Research, Wellington, 1961).
24. Sackett, W. M., Poag, C. W. & Eadie, B. J. *Science* **185**, 1045–1047 (1974).

α -Particle radioactivity of hot particles from the Esk estuary

E. I. Hamilton

Institute for Marine Environmental Research, Prospect Place,
The Hoe, Plymouth PL1 3DH, UK

Transuranium radionuclides (Pu, Am and Cm) present in effluents discharged into the north-east Irish Sea by British Nuclear Fuels Limited, Windscale, Cumbria, UK, are found in sediment and biota of the Esk estuary ~10 km to the south. The site of the present investigation was at Newbiggin and the materials examined were suspended particulate debris samples at the sea surface, bottom sediments and some forms of biota collected in September 1977. It is shown here that hot particles (defined as small volumes of material emitting α particles recorded in a dielectric detector as dense clusters of tracks from a common origin) found in the estuary are likely to be original effluent debris derived from the processing of Magnox uranium fuel elements and not formed *in situ* as a result of natural processes common to the estuary.

Methods of analysis described elsewhere^{1–4} use a new dielectric α -track detector, CR-39 (ref. 5), in association with conventional chemical separation techniques and α -particle spectrometry using surface barrier detectors. The CR-39 method involves placing a sample in contact with a sheet of detector and after a suitable period of exposure the α particles recorded are developed by etching at 80 °C in 6.0 M NaOH for 3 h. After etching, the detector is viewed with the naked eye and α tracks associated with hot particles can be easily observed against a background of randomly distributed tracks; analysis of α -particle track distribution is then determined using a conventional light microscope with $\times 5$ or $\times 16$ objectives and $\times 10$ oculars; quantitative analysis of track densities is determined using a Quantimet image analyser.

In the Esk estuary (at the Newbiggin rail viaduct), after slack water and when the flood tide is just starting, the water becomes very turbid³; for a short period of time (15–45 min depending on the energy of the tide), fine-grained (<100 μ m) bottom deposited silts become resuspended into the water column (0.1–0.5 m deep) and just reach the surface, although in calm water conditions the top ~1 mm of water is often clear. In calm conditions the paper side of a sheet (30 $\times$ 30 cm) of Whatman

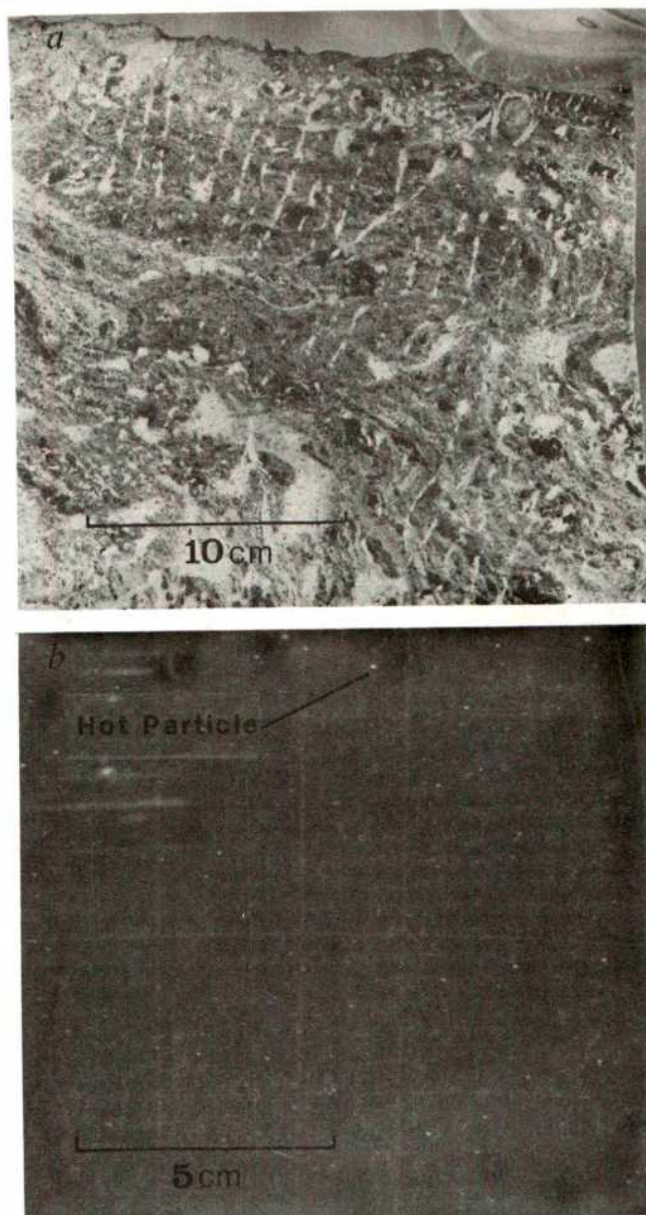


Fig. 1 Identification of the distribution of hot particles in sediment using CR-39 and surface barrier α spectrometry. *a*, Seawater surface sediment absorbed onto Whatman Benchkote. *b*, Etched CR-39 illustrating (white spots) sites of hot particles in *a*. *c*, Thick source α spectra of a single hot particle showing the presence of U, Pu and Am radionuclides.

Table 1 $^{239+240}\text{Pu}$: ^{238}Pu activity ratios, percent abundance of $^{238+239+240}\text{Pu}$ and ^{241}Am in various materials from the Esk estuary, Cumbria, UK and adjacent areas collected in 1977

Material	α particle activity ratio $^{239+240}\text{Pu}/^{238}\text{Pu}$	α emitting radionuclides (% abundance)		Mean Pu/Am (α activities)	Ref.
		$^{238+239+240}\text{Pu}$	^{241}Am		
Newbiggin silts ($n=8$)	3.6 ± 0.4	46.3	53.7	0.9 ± 0.2	3
Top 2 cm sediment core 1176 Newbiggin	3.4 ± 0.4	57.4	42.6	1.4 ± 0.4	3
Regional sediments ($n=40$)	4.3 ± 0.5	53.0	47.0	1.1 ± 0.3	21
Benthic animals ($n=56$)	4.2 ± 0.3	46.2	53.8	0.9 ± 0.5	3
Fish ($n=5$)	3.9 ± 0.5	54.7	45.3	1.2 ± 0.2	21
Mean	3.9 ± 0.3	—	—	1.1 ± 0.2	
<i>Fucus vesiculosus</i>					
Ravenglass	3.7	86.7	13.3	6.5	3
Newbiggin	3.8	65.8	34.2	1.9	21

Data obtained following chemical separation and surface barrier α spectrometry.

Benchkote is touched onto the sea surface, just penetrating the clear surface layer of water if present, to which particulates recently resuspended into the water column adhere and are removed. Visible swirls of particulate debris present on the sea surface are retained by this technique as shown in Fig. 1a. Recently exposed bottom sediments are sampled similarly and if no pressure is applied to the paper, only the top layer of sediment (~ 1 mm) is removed. The Benchkote is then allowed to dry until just damp when it is placed against a sheet of CR-39 and exposed for between 2 h and 10 days. After the detector is etched two features are visible to the naked eye (Fig. 1b); first, the presence of diffuse opalescent markings which reproduce the macrostructure of the adhered sediment, and second, randomly distributed white spots ($<100 \mu\text{m}$ mean particle diameter) which represent the intense activity originating from hot spots. The position of the hot particles in the detector is noted and the corresponding site on the Benchkote identified. A small area of Benchkote enclosing a hot particle is then cut out; typically the level of diffuse activity associated with the area of paper removed is insignificant and all the activity originates from the hot particle. Briefly the following analytical techniques are then used. (1) The total number of α tracks from each separated particle is determined using the Quantimet analyser calibrated for α -particle detection efficiency against a standardized zinc sulphide screen counter⁴. When the track density is too intense with the result that individual tracks cannot be recognized, the total activity of the particle is determined by the zinc sulphide screen method; typically particle activities are between 0.1 and 3 pCi for total α emissions. (2) The hot particle attached to the Benchkote is then mounted in an energy calibrated surface barrier α spectrometer and the α spectrum determined. In the samples investigated so far ($n=28$), four of the spectra were partially degraded; these particles were dissolved in hot hydrochloric acid together with ^{236}Pu and ^{243}Am yield monitors, the detrital Benchkote removed and the leachate, after adjustment to pH of a plating solution, electroplated onto a polished stainless steel planchet. The total α activity recorded after electroplating agreed with that obtained by the zinc sulphide method and validates hydrochloric acid being used instead of nitric acid³ as a leaching agent to reduce the degradation of the Benchkote filter paper. None of the particles examined gave a totally degraded α spectrum and using the procedures described by Evans⁶ individual energies were identified. α -particle composition of the particles varied considerably; Pu and Am were always present, curium radionuclides were not always present; in six samples low energy emissions at ~ 4.8 meV and 4.2 meV were

observed and are tentatively identified as originating from $^{234}\text{U} + ^{237}\text{Np}$ and ^{238}U respectively.

When electroplating techniques are used without prior chemical separation of Pu and Am, because of energy overlap the radionuclides determined are $^{239+240}\text{Pu}$, $^{238}\text{Pu} + ^{241}\text{Am}$, $^{243+244}\text{Cm}$ and ^{242}Cm ; compared with standard chemical techniques³ no separation of ^{238}Pu from ^{241}Am is obtained. To convert the data into conventional terms the ^{238}Pu component had to be identified. Using chemical separation methods, together with published data, a $^{238}\text{Pu}/^{239}\text{Pu}$ ratio of 3.9 ± 0.3 determined on a large number of sediments and biota collected in September 1977 from the area, was used to determine the ^{238}Pu component of the $^{238}\text{Pu} + ^{241}\text{Am}$ α -spectra peaks. The activity of individual α emitters from Pu and Am was then calculated in terms of the percentage abundance for total Pu (α) and Am; the data are given in Table 1 and Fig. 2 and show that although the mean Pu/Am ratio for sediments and biota of the area is 1.1 ± 0.2 , the sea surface particulates showed a much greater range (Pu/Am, 3.2 ± 3.0) with a distinct trend towards an enrichment in Pu; none of the samples was significantly enriched in Am. A preliminary examination of hot particles ($n=10$) from the top 10 cm of deposited sediments gave a Pu/Am ratio of 0.9 ± 0.2 ; however, the morphology of these particles differs from those collected near the sea surface and will be discussed elsewhere. Note, however, that there are fewer hot particles present in buried sediments, and the specific activity is less. Data for *Fucus vesiculosus* indicates some degree of Pu enrichment which may be related to the entrainment of particles on fronds³.

In the absence of any quantitative data which describes the isotopic composition of transuranics, together with their chemical and physical form in fresh effluent released by BNFL, they cannot be identified in the Esk. However, circumstantial evidence allows some interpretation of the data. Using data presented by Nair⁷ it is calculated that the specific α radioactivity of aged Magnox fuel is $\sim 0.2 \text{ Ci g}^{-1}$. Measurement of particle size, using a technique based on the diameter of hotspots⁴ shows that the highest activity is usually associated with particles with a mean particle diameter of $<20 \mu\text{m}$; assuming that the density of these particles is ~ 10 they would have a maximum total α activity of $1\text{--}10 \mu\text{Ci g}^{-1}$, and therefore they do not consist of unprocessed uranium fuel. In the Esk estuary undisturbed sediments contain significant levels of short-lived radionuclides such as $^{95}\text{Zr-Nb}$ and ^{242}Cm , hence seaborne transit times for effluent from the outfall to the Esk are rapid enough for short-lived radionuclides to reach the area; indeed the estimated transit time is $<1\text{--}3$ months⁸, hence the presence or absence of water transported particulate debris at any single sampling time will depend on the previous weather conditions

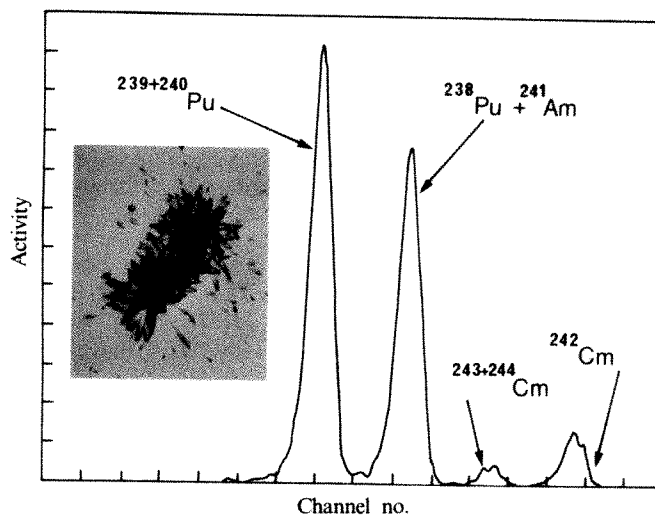


Fig. 2 Photograph and direct α spectra for a hot particle containing Pu, Am and Cm radionuclides sampled at the sea surface.

Table 2 Mean monthly discharges (Ci yr^{-1}) for Pu and Am α radioactivity) from the uranium reprocessing plant together with Pu/Am ratios⁹

Year	$^{238}\text{Pu} + ^{239+240}\text{Pu}$	^{241}Am	Pu/Am
1977	984	96	10.3
1976	1272	324	3.9
1975	1200	984	1.2
1974	1248	3192	0.4

common to the region. Table 2 provides data for the Pu/Am ratio, derived from mean monthly discharge data which, in 1977⁹ was 10.3; in previous years it was significantly lower. The 1977 Pu/Am ratio for Esk sediments and biota (see Table 1 and Fig. 3) can be matched with 1975 releases, thus requiring a delay of ~2 yr before they become characteristic for the Esk. No data are available for the monthly discharges of Pu or Am released in the effluent, but from data presented by Baxter¹⁰ it seems probable that maximum releases occur during September–December of each year. Other data³ show that it was not until 1979 that sediment and biota of the Esk estuary showed enhanced levels of Pu, supporting the concept of a delay period ~2 yr before 1977 effluents became dominant in the region. Recently¹¹ it has been suggested, on the basis of derived sedimentation rates for Esk sediments, that there is a finite time lag of 2–3 yr between the plutonium discharges from BNFL and the arrival of plutonium in these sediments. Fundamental differences exist in the determination of the apparent time lag and are discussed elsewhere¹². Hence on this basis surface sediment of the Esk in 1977 would not be expected to reflect isotope ratios for current releases; through rapid transport of suspended particulates in the water column such debris could be present but its contribution to bulk deposited surface sediments would be small and difficult to identify. Note that, compared with sediments, uptake of Pu and Am by biota would be expected to reflect isotope ratios in seawater common to an area when the samples were taken.¹³

Six of the spectra obtained from the hot particles showed the presence of low energy emissions at 4.2 MeV and 4.8 MeV; for other samples the energy cutoff below the $^{239+240}\text{Pu}$ peak was sharp. These energies could arise from $^{234}\text{U} + ^{237}\text{Np}$ and ^{238}U . In 1978, 16 Ci of ^{237}Np was discharged by BNFL¹⁴ (no data are available for 1977) which should be detectable as it is similar to the discharge level for ^{242}Cm (15 Ci); it is assumed that ^{237}Np has always been a component of the BNFL effluent. Recent measurements confirm the presence of ^{237}Np in Esk silts at levels comparable with those found for Cm radionuclides. ^{237}Np in *Porphyra* from Sellafield, Cumbria is reported¹⁶ at a similar level to the ^{242}Cm abundance. Use of the delayed neutron method for uranium assay has shown² that in spite of a discharge of considerable quantities of uranium in BNFL effluent (for example, 10,936 kg in 1977)¹⁴ the levels for uranium in the Esk sediments are normal. Recent measurements¹⁵ have shown that the isotopic composition of uranium in Esk sediments does not show the presence of significant levels of depleted uranium despite problems associated with the production of ^{234}U from the decay of ^{238}Pu . However, the hot particles from surface water of the Esk seem to contain α emitters which would be compatible with the presence of U and ^{237}Np . The spectra do not contain ^{226}Ra or daughter products or thorium isotopes which could be attributed to natural sources. Therefore, the presence of the (U–Np)–Pu–Am–Cm radionuclides in some hot particles is interpreted as evidence that they represent original effluent materials. The lifetime for uranium-bearing particles in the sea is not known. A 4×0.5 mm uranium rod placed in seawater (oxygenated or reduced) decomposes within a few weeks to give dissolved species and finely disseminated uranium oxides which then slowly dissolve. It seems unlikely that any uranium-rich particle would survive intact for very long in the sea, thus they are unlikely to be present in deposited sediments, but could occur in the water column as a result of recent transport.

Uranium behaves in a conservative manner in seawater and using levels for ^{137}Cs (also conservative) in seawater off Windscale as a measure of effluent levels, it is calculated that if the 1977 discharges of U were totally dissolved in the sea this would give rise to a concentration of $0.03 \mu\text{g U l}^{-1}$ for offshore water, and $0.1 \mu\text{g U l}^{-1}$ for near-shore water, which would be difficult to detect in the presence of normal levels of $\sim 3 \mu\text{g U l}^{-1}$ for seawater².

Spent Magnox fuel rods are usually kept submerged in tanks held at pH 11.5, containing only trace quantities of Cl^- , CO_3^{2-} and SO_4^{2-} ions to prevent solution of the Magnox; usually imperfection and impurities cause some uranium fuel loss and this accumulates on the bottom of the tanks. Before recent improvements in waste treatment, the final low activity tanks were maintained at pH 2–3 which would precipitate iron, present as a contaminant, and coprecipitate the actinides¹⁷. Before discharge into the sea the effluent is treated with ammonium hydroxide which will predispose any remaining actinides in solution to assume less soluble forms (that is, colloids). Magnesium derived from the Magnox cladding is also likely to be present in the effluent. The released effluent has a complex composition and apparently contains particulate phases, some of which would dissolve in the sea (when dispersed) according to individual compositions. Radionuclide compositions of particles found in the Esk are similar to those of the original effluent before discharge, therefore are highly unlikely to have been formed *in situ* in the Esk. In the sea, magnesium hydroxide particles are likely to dissolve slowly while labelled iron particles would be transported further, and then become subject to solution processes in the organic-rich surface sediments of the Esk as a result of geochemical recycling. In the Esk interactions with freshwater and flocculation processes which occur at the salt–freshwater interface would assist dissolution of particles.

Dissolution of hot particles is supported by the form of the particulates. Hot particles with well defined form such as shown in Fig. 2 have been found in surface sediment and water, but are not always associated with the low-energy α emitters. In buried sediments occasional hot particles are present, but usually of lower activity. They often tend to have dimensions of ~ 10 (a and b axis) $\times 30$ (c axis) μm and would not be inconsistent with

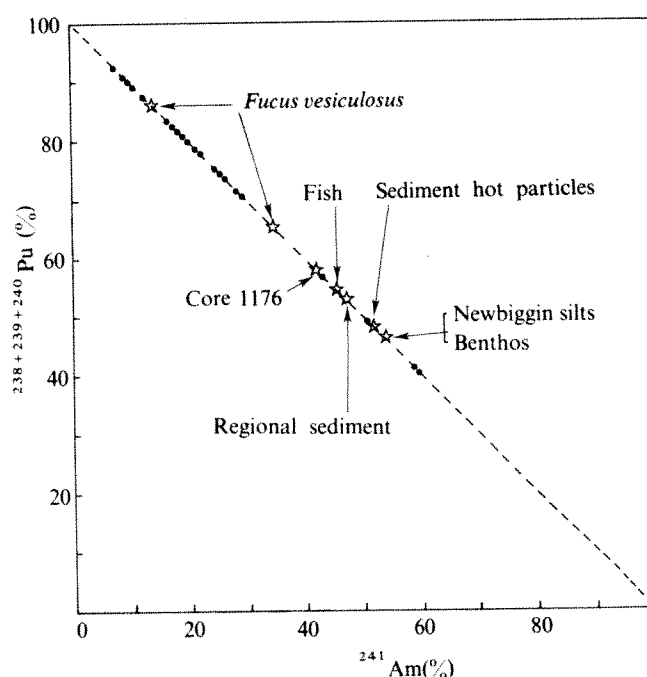


Fig. 3 Ratio Pu/Am for sea surface particulates and other material (see Table 1).

the gradual solution of the surface water particles shown in Fig. 2. The hot particles associated with particulate matter sampled at the sea surface, and hence available for transfer into marine aerosols, have a mean particle diameter suitable for entry into the lung of man. In relation to potential health hazards to man the levels of currently released radionuclide effluents may be of more concern than those retained in older sediments where hot particles are less abundant and a diffuse distribution of α activity is more common, possibly reflecting redistribution as a result of geochemical recycling. Apart from the release of transuranics in effluent into the sea there is also a need to consider releases from the BNFL stacks and from the uncovered uranium fuel rod cooling ponds. Cambray and Eakins¹⁸ have shown that Pu is present in soils up to ~10 km from the Cumbrian coast and is either derived from direct emissions from Windscale or indirectly from the sea. Gorham¹⁹ has shown that within the region, up to a distance of at least 51 km from the coast, almost all the sodium and chloride ions in tarns (altitude 15–275 m) is derived from sea spray. Cambray and Eakins noted a correlation between Na and Pu and favour the sea surface as a major source for the Pu. Apart from the transport of conservative species of Pu and Am in droplets of seawater in marine aerosols the possibility exists for entrainment of the small sea surface hot particles which I have also found in soils, lichens and one ombotrophic peat inland from the coast. So far their distribution has not been evaluated because of difficulties in developing an acceptable programme of valid statistical sampling of such diffuse distributions.

Using conventional air sampling methods Pattenden *et al.*²⁰ note that Pu isotopes, ²⁴¹Am and ¹³⁷Cs are present in airborne and deposited material in west Cumbria in excess of the quantities expected for nuclear weapon tests. Excess actinides are considered to be derived from sea-spray and surf, but based on measurements of ¹³⁷Cs this cannot be accounted for by transfer of bulk seawater in the form of droplets. They note that conventional sampling methods are not efficient for samples with a mean particle diameter of >20 μ m because the inertia of large particles limits an ability to follow a moving air-stream. Using the CR-39 method I have observed that while the highest specific activity of resuspended air and sand particles is usually associated with those of a mean particle diameter of <20 μ m, there is an abundance of uniformly-labelled large particles which consist of coated sand grains (mean particle diameter 100–300 μ m) and ferruginous flocs¹⁵. In an examination of ~1,000 particles from the Esk region (resuspended silts, rocks, lichens) only one ~500 μ m particle was found to have a specific activity comparable with the sea-surface hot spots. The contribution of diffuse labelled large particles to sea-surface type hot particles is being investigated and is likely to be significant. CR-39 is capable of identifying the distribution of the hot particles on foliage and the surface of soils distant from Windscale and such a programme has been implemented. Because of the considerable uncertainties in even measuring maximum permissible levels of Pu in the human lung it would be useful to determine the distribution and composition of these hot particles in Cumbria and adjacent areas and evaluate their importance to man.

I thank R. J. Clifton and P. J. Moore for useful discussions. The Institute for Marine Environmental Research is a component of the Natural Environment Research Council.

Received 24 September 1980; accepted 10 February 1981.

1. Hamilton, E. I. *Symp. on Detection of Radiation in Environmental and Biological Materials* (CEGB, London, 1978).
2. Hamilton, E. I. *Mar. Ecol. Prog. Ser.* **2**, 61–73 (1980).
3. Hamilton, E. I. & Clifton, R. J. *Mar. Ecol. Prog. Ser.* **3**, 267–277 (1980).
4. Hamilton, E. I. & Clifton, R. J. *Int. J. Appl. Isotopes Radiat.* (in the press).
5. Cartwright, B. G., Shirk, E. K. & Price, P. B. *Nucl. Instrum. Meth.* **153**, 457–460 (1978).
6. Evans, R. D. A. *Prog. Rep. AT (30-1)-952-MIT-952-4*, 132–162 (Radioactivity Center, MIT 1967).
7. Nair, S. U.K. C.E.G.B. Rep. RD/B/N4349, 1–34 (1978).
8. Hetherington, J. A. & Jefferies, D. F. *Neth. J. Sea. Res.* **8**, 319 (1974).
9. Pentreath, R. J., Lovett, M. B., Harvey, B. R. & Ibbett, R. D. *Biological Implications of Radionuclides Released from Nuclear Industries Vol. II* (IAEA, Vienna, 1979).
10. Baxter, M. S., McKinley, I. G. M., Mackenzie, A. B. & Jack, W. *Mar. Pollut. Bull.* **10**, 116–120 (1979).

11. Aston, S. R. & Stanner, D. A. *Nature* **289**, 581–582 (1981).
12. Clifton, R. J. & Hamilton, E. I. *Est. Coast. Shelf Sci.* (submitted).
13. Hodge, V. F., Koide, M. & Goldberg, E. D. *Nature* **227**, 206–209 (1979).
14. A. Rep. *Radiation Discharges and Monitoring of the Environment 1978* (British Nuclear Fuels, London 1979).
15. Hamilton, E. I., Clifton, R. J. & Stevens, H. (in preparation).
16. Hampson, B. L. & Tennant, D. *Analyst* **98**, 873–885 (1973).
17. Hetherington, J. A., Jefferies, D. F. & Lovett, M. B. In *Impacts of Nuclear Releases into the Aquatic Environment*, 210 (IAEA, Vienna, 1975).
18. Cambray, R. S. *UKAEA AERE Rep. R-9807*, HL80/2061 (C10) 1–15 (HMSO, London, 1980).
19. Gorham, E. *Phil. Trans. R. Soc. B*, 147–178 (1958).
20. Pattenden, N. J., Cambray, R. S., Playford, K., Eakins, J. D. & Fisher, E. M. R. *UKAEA AERE Rep. R-9857*, HL80/2746 (C10) 1–22 (HMSO, London, 1980).
21. Hunt, G. J. *Aquatic. Envir. Mon. Rep. MAFF—FRL No. 3*, 1–36 (1979).

Specific acyclic isoprenoids as biological markers of methanogenic bacteria in marine sediments

S. C. Brassell, A. M. K. Wardroper, I. D. Thomson, J. R. Maxwell & G. Eglinton

Organic Geochemistry Unit, University of Bristol, School of Chemistry, Cantock's Close, Bristol BS8 1TS, UK

The widespread occurrence of extended hopanoids in sediments and petroleum illustrates the importance of bacterial lipid contributions to geological materials¹. In archaeobacteria, however, hopanoids are absent¹; their role as structural components of biomembranes is fulfilled by acyclic isoprenoids^{1–5}. Recent studies^{2–6} of the lipid constituents of archaeobacteria have greatly extended the range of acyclic isoprenoid skeletons known in organisms (Fig. 1). In particular, isoprenoids with head-to-head linkages have been identified^{3–5}, and such compounds (for example, 3,7,11,15,18,22,26,30-octamethyldotriacontane, I) have been recognized in petroleum^{7,8} and as degradation products of Messel shale kerogen^{9–11}. Here we report the first recognition of 2,6,10,15,19-pentamethyleicosane (II), a known component of methanogens^{5,6}, in marine sediments of Recent to Cretaceous age (Table 1) and suggest that it and certain other acyclic isoprenoids may be used as biological markers for methanogens.

Until recently, the geologically occurring isoprenoid alkane, squalane (III), first characterized in a Nigerian oil¹², was generally believed to originate from geochemical reduction of squalene (IV). The discovery of squalane in several immature marine sediments^{13–17}, including Recent sediments^{13,14}, suggested that such a reduction must occur rapidly in the upper layers of sediment and therefore be microbially mediated. An alternative explanation is that squalane represents a direct input from organisms and the recent recognition of squalane in archaeobacteria^{5,6} supports this proposition. If, in general, archaeobacteria are important contributors of lipids to sediments, one would expect to find other compounds uniquely biosynthesized by these organisms in geological materials. One such compound seems to be 2,6,10,15,19-pentamethyleicosane, which has been found in methanogens^{5,6} but not in the other types of archaeobacteria, or in marine phytoplankton, such as diatoms, dinoflagellates, silicoflagellates and coccolithophores, which are major sources of sediment lipids. In addition, it has not been detected in zooplankton, although these organisms frequently contain another acyclic isoprenoid alkane, pristane, derived from chlorophyll¹⁸.

2,6,10,15,19-Pentamethyleicosane (II), squalane (III) and lycopane (V) were identified in the sediments by comparison of their mass spectra, obtained by computerized gas chromatography-mass spectrometric¹⁹ analysis of hydrocarbon fractions, with literature and/or reference spectra. Each sediment was extracted using ultrasonication; the total lipid extract was separated by TLC yielding a hydrocarbon fraction which, in some cases, was further separated by urea adduction. The mass

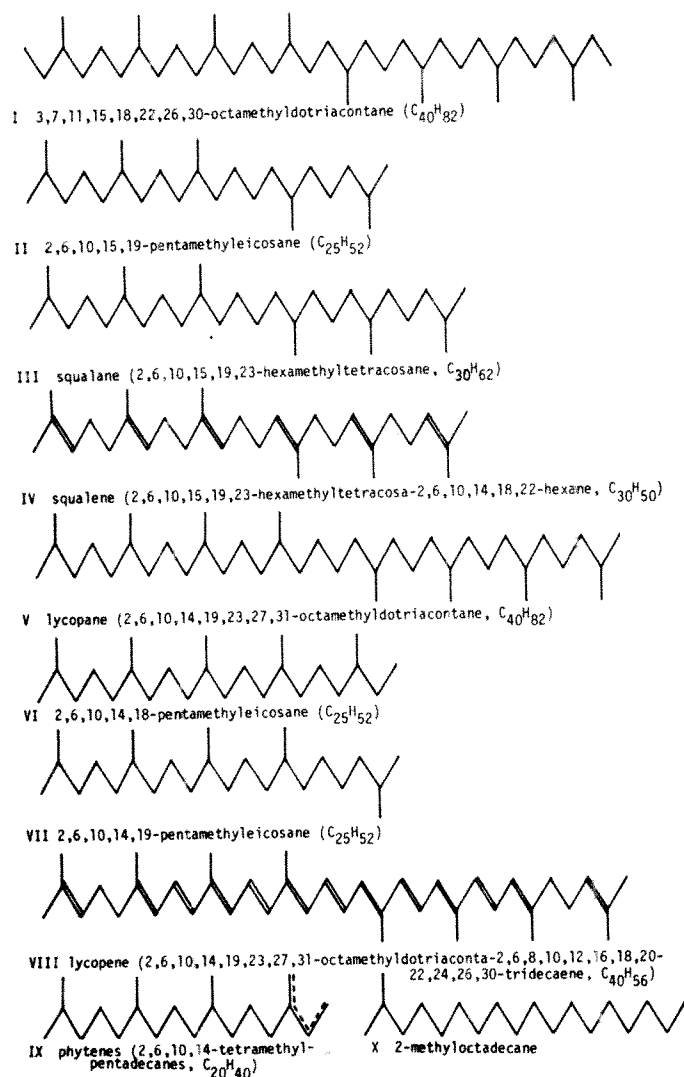
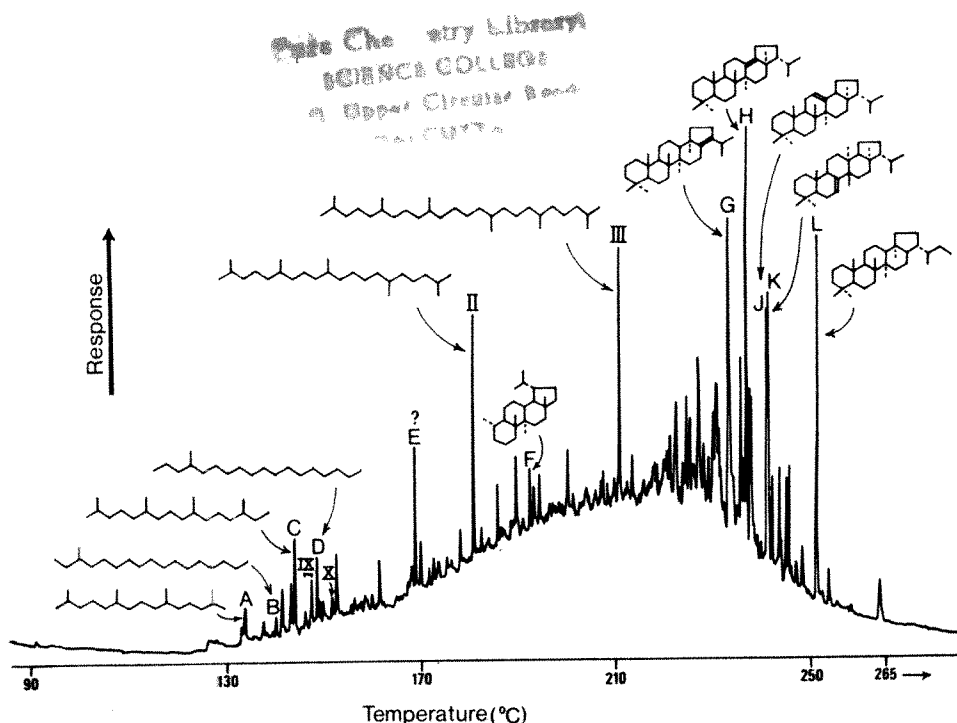


Fig. 1 Structures of acyclic hydrocarbons.

Fig. 2 Gas chromatogram (GC) of the branched/cyclic hydrocarbons of a Pleistocene diatomaceous clay/silty claystone (DSDP Leg 57, sample 57-440B-8-4; Table 1) from the Japan Trench. Peaks: A, pristane; B, 3-methylheptadecane; C, phytane; IX, phytanes; D, 4-methyloctadecane; X, 2-methyloctadecane; E, $C_{25}H_{44}$ bicyclic alkadiene of unknown structure^{44,45}; II, 2,6,10,15,19-pentamethyleicosane; F, tetracyclic alkane related to lupane^{46,47}; III, squalane; G, hop-17(21)-ene; H, neohop-13(18)-ene; J, neohop-12-ene; K, fern-7-ene; L, 17 β H,21 β H-homohopane. GC conditions: 20 m OV-1 wall-coated glass capillary, programmed from 50 to 265 °C at 4 °C per min after splitless injection at ambient temperature. A prominent feature of the GC trace is the apparent 'hump' which in crude oils and polluted sediments would be attributed to a background of thermally matured hydrocarbons. In this sample, however, the concentrations of other individual hydrocarbons (for example, steranes and hopanes) commonly associated with the 'hump' is small; hence, it is uncertain whether it represents highly biodegraded thermally matured or natural inputs or, alternatively, shipboard contamination with drilling lubricant.



spectrum of 2,6,10,15,19-pentamethyleicosane (II) is distinct from those of VI (2,6,10,14,18-pentamethyleicosane) and VII (2,6,10,14,19-pentamethyleicosane) for several reasons^{6,16,17}, particularly the absence of m/z 253 in II. The assignments of squalane (III) and lycopane (V) were corroborated in a number of cases by gas chromatographic co-injection with authentic standards. Components were quantified from their gas chromatographic responses.

The recognition of 2,6,10,15,19-pentamethyleicosane (II) in three marine environments where methanogens have been recognized and/or high concentrations of methane have been found (Table 1) suggests the working hypothesis that this alkane marks a contribution of lipids from methanogens to marine sediments. An isomer, 2,6,10,14,18-pentamethyleicosane (VI), recently found in thermoacidophilic bacteria⁶, has previously been reported as a constituent of various shales²⁰, oils^{7,8,21-23} and sediments²⁴. The occurrence of squalane (III) may also reflect methanogenic activity in these marine sediments (Table 1), although this compound has been found in both thermoacidophilic and methanogenic bacteria⁶. However, significant inputs to marine sediments from thermoacidophilic and halophilic bacteria are improbable, although such archaeobacteria may be important sources of lipids in other environments, for example in the Dead Sea²⁵. Lycopane (V) is another isoprenoid alkane found in significant quantities (see Table 1) in Recent^{14,16,17} and ancient²⁶ marine sediments. It could be formed by diagenetic reduction (perhaps microbial) of lycopene (VIII)^{14,16,17} but, by analogy with squalane, a direct bacterial origin is also possible. Indeed, the Recent sediments in which lycopane has been reported are all environments where methanogenic activity has been documented (Table 1).

In older and more deeply buried sediments, after methanogenic biological activity has ceased, the presence of methane and the lipid signatures may attest its former existence. Hence, in the Pleistocene to Miocene sediments of the Cariaco and Japan Trenches, the presence of methane in high concentrations²⁷⁻²⁹ and the recognition of the isoprenoid alkanes II and III (Table 1) indicate past biological methanogenesis in these sediments. In more mature sediments the presence of methane alone cannot be used as an indicator of former methanogenic biological activity. For example, the methane in the Moroccan Basin turbidites (Table 1) is thought to be migrated rather than

Table 1 Concentrations of 2,6,10,15,19-pentamethyleicosane (II), squalene (III) and lycopane (V) in selected marine sediments

Location (lat, long)	Sub-bottom depth(m)	Geological age (Myr BP)	Dominant lithology	Concentrations (ng per g dry sediment weight)			
				II	III	V	CH ₄
Walvis Bay (22°35'S, 13°45'E) 1/75 (0-0.1)	0.05	Holocene (<0.01)	Diatomaceous ooze	250	ND	180	High*
Gulf of California (27°15'N, 111°31'W) 64-481-2-2	7.5	Holocene (<0.01)	Diatomaceous ooze	16	23	68	High†
Cariaco Trench (10°43'N, 65°11'W) 15-147B-1-2 15-147B-5-5 15-147C-3-3	3 48 138	Holocene (<0.01) Pleistocene (~0.1) Pleistocene (~0.25)	Calcareous clay Calcareous clay Calcareous clay	1.1 + 6.7	0.7 + 5.3	14 + 57	} High‡
Japan Trench (39°44'N, 143°48'E) 56-435-13-2 (39°56'N, 145°33'E) 56-436-11-4 (39°44'N, 143°56'E) 57-440-5-6 57-440A-7-6 57-440B-3-5 57-440B-8-4 57-440B-23-4 57-440B-39-3 57-440B-68-2	115 99 42.5 139 166 212 354 504 778	M. Pleistocene (~3) U. Pliocene (~2) U. Pleistocene (~0.5) L. Pleistocene (~1) L. Pleistocene (~1) L. Pleistocene (~1) U. Pliocene (~2) U. Pliocene (~4) U. Miocene (~6)	Diatomaceous clay Diatomaceous vitric clay Diatomaceous silty clay Diatomaceous silty clay Diatomaceous clay Diatomaceous clay/silty claystone Diatomaceous claystone/silty claystone Diatomaceous claystone Silty claystone	0.8 0.3 0.3 1.5 1.9 1.7 0.5 2.8 1.4	0.4 0.6 0.8 2.8 3.1 1.7 0.4 1.8 1.7	ND ND ND ND ND ND ND ND ND	446§ 5 115 842 803 1596 151 215 289
Blake-Bahama Basin (28°14'N, 75°37'W) 44-391A-13-0 44-391C-6-3 44-391C-12-4	525 691 959	L. Miocene (~20) U. Albion (~100) Aptian (~110)	Siliceous mudstone Claystone Calcareous claystone	+ + ND	ND + ND	ND ND ND	} NA
Moroccan Basin (32°50'N, 10°48'W) 50-416A-9-2 50-416A-14-4 50-416A-18-2 50-416A-23-3 50-416A-28-4	1179 1228 1266 1313 1362	Hauterivian (~125) Valanginian (~130) Valanginian (~130) Valanginian (~130) Valanginian (~130)	Turbidite Turbidite Turbidite Turbidite Turbidite	0.4 0.3 0.1 Tr Tr	0.5 0.5 Tr 0.2 0.1	ND ND ND ND ND	11.4 ND ND 14.9 ND

Detailed analyses of the lipid distributions of II, III and V are reported elsewhere^{15-17,36-38}. ND, not detected; NA, not analysed; +, minor amounts not quantified; Tr, trace (<0.1 ng per g) amounts of component present.

*CH₄ found (90-2,926 and 10-91 ng per g) in subsections of sediment cores up to 2.3-m depth from nearby locations (22°47'S, 14°26'E and 23°14'S, 14°01'E, respectively) and attributed to methanogenic bacterial activity²⁷.

†CH₄ occurs in Site 481 sediments³⁹ and is abundant (2,190-7,520 ng per g) in the subsurface sediments (2.51-3.45 m) collected about 8 km to the north-east (27°23'N, 111°27'W), with $\delta^{13}\text{C}$ values (-68.8 to -73.5) indicating its biogenic origin⁴⁰. Methanogenic bacterial activity has been recognized in an overlying sample (64-481-1-2) from 2.9-m depth⁴¹ and shallow sediments (0.5-1.7 m) from a nearby site (27°09'N, 111°08'W)⁴².

‡CH₄ detected in gas pockets to 175-m depth at Site (147 (ref. 28) and found in high concentrations (>200 ng per g) in other shallow sediments (0-1.6 m) from the area²⁷. Methanogenic bacteria have been isolated from nearby surface sediments within the western basin of the Trench (10°37'N, 65°26'W)⁴³. Lycopane has also been recognized in shallow sediments (0-5.5 m) from the eastern basin (10°32'N, 64°39'W)¹⁴.

§Data for adjacent samples from ref. 29.

||Data for adjacent samples from ref. 30.

indigenous gas³⁰. Hence, the presence of the acyclic isoprenoid alkanes provides a better geological record of biological methanogenesis than the presence of methane.

In the Holocene sediments the differences in the relative proportions of II, III and V may result from dissimilar methanogen populations, as methanogen lipid compositions are species dependent^{5,6}. In addition, the variability in their concentrations in these sediments may reflect the intensity of biological methanogenesis. The deeper and older sediments tend to contain lower concentrations than do the Holocene sediments (Table 1), partly because other microorganisms may have degraded the methanogenic lipids. Other factors may be involved, for example, the known flux of terrestrial material in the Japan Trench sediments and Moroccan Basin turbidites may have diluted the methanogenic lipid signature³¹. Such factors, however, tend to restrict rather than invalidate the use of acyclic isoprenoid alkanes as biological markers of methanogens. Although the Cretaceous samples contain only low concentrations of II and III, the mere presence of these compounds is presumptive evidence for methanogenesis at the time of deposition. Methanogenic bacteria seem, therefore, to be important contributors of lipids to deep-sea sediments at present and in the geological past, presumably as a result of their activity at appropriate E_h (redox potential) zones in the sub-bottom sediments. The significance of archaeobacterial contri-

butions to petroleum has already been demonstrated^{7,8}, but present evidence suggests that specific acyclic isoprenoid alkanes may also be environmental indicators. That they are important constituents of sediments can be seen from their relative abundances in branched/cyclic (urea non-adducted) hydrocarbon fractions (Fig. 2). Future investigations should include stereochemical determination of these acyclic isoprenoid alkanes in sediment extracts, as their configurations will reflect their origin and diagenetic history^{32,33}, together with studies of the stereochemistry of these compounds in methanogenic bacteria.

Other isoprenoid hydrocarbons may also arise from methanogens. Squalene (IV) and phytanes (IX) have been found in young marine sediments^{16,17}, but could originate from other sources—for example, phytanes may be products of phytol or dihydrophytol diagenesis³⁴. Thus, these compounds cannot generally be regarded as definitive indicators of archaeobacterial inputs to sediments. An additional group of acyclic hydrocarbons biosynthesized by archaeobacteria⁶ is branched alkanes (for example, 2-methyloctadecane, X); these compounds also occur in deep-sea sediments¹⁷ (Fig. 2). Indeed, most of the prominent branched/cyclic hydrocarbons of the Japan Trench sediments (Fig. 2) can be ascribed to bacterial origins. In view of their recent recognition in a bacterium³⁵, the presence of hop-17(21)-ene (G), neohop-13(18)-ene (H) and fern-7-ene (K) in

immature sediments probably reflects autochthonous inputs from bacteria rather than terrestrial inputs from ferns¹⁵⁻¹⁷. It seems likely, by analogy, that neohop-12-ene (J) is also of bacterial origin.

Thus, there is growing evidence in support of major bacterial lipid inputs to marine sediments. The presence of components which seem to be biosynthesized uniquely by methanogens, such as 2,6,10,15,19-pentamethyleicosane, suggests that these bacteria are important contributors of alkanes to marine sediments.

We thank the NERC (GR3/2951 and GR3/3758) and NASA (subcontract from NGL-05-003-003) for funding. A.M.K.W. and I.D.T. acknowledge studentships from the SRC, S.C.B. from the NERC. We thank Dr R. S. Oremland for information about methanogens in Gulf of California sediments, Mrs A.P. Gower for technical assistance, and colleagues at Bristol for access to extracts^{15,19}. Sediment samples were provided by the Deep Sea Drilling Project (JOIDES), with the exception of the Walvis Bay sediment, which was kindly supplied by Dr R. J. Morris.

Received 11 August 1980; accepted 4 February 1981.

1. Ourisson, G., Albrecht, P. & Rohmer, M. *Pure appl. Chem.* **61**, 709-729 (1979).
2. Kates, M. in *Ether Lipids, Chemistry and Biology* (ed. Synder, F. S.) 351-398 (Academic, New York, 1972).
3. Tornabene, T. G. & Langworthy, T. A. *Science* **203**, 51-53 (1979).
4. De Rosa, M., De Rosa, S., Gambacorta, A. & Bu'Lock, J. D. *Chem. Commun.* 514-515 (1977).
5. Tornabene, T. G., Langworthy, T. A., Holzer, G. & Oro, J. *J. molec. Evol.* **13**, 73-83 (1979).
6. Holzer, G., Oro, J. & Tornabene, T. G. *J. Chromatogr.* **186**, 795-809 (1979).
7. Moldowan, J. M. & Seifert, W. K. *Science* **204**, 169-171 (1979).
8. Albaiges, J. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 19-28 (Pergamon, Oxford, 1980).
9. Michaelis, W. & Albrecht, P. *Naturwissenschaften* **66**, 420-422 (1979).
10. Chappe, B., Michaelis, W. & Albrecht, P. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 265-274 (Pergamon, Oxford, 1980).
11. Chappe, B., Michaelis, W., Albrecht, P. & Ourisson, G. *Naturwissenschaften* **66**, 522-523 (1979).
12. Gardner, P. M. & Whitehead, E. V. *Geochim. cosmochim. Acta* **36**, 259-263 (1972).
13. Dastillung, M., Albrecht, P., Tissier, M. J. in *Géochimie Organique des Sédiments Marins Profonds. Orgon I; Mer de Norvège*, 209-228 (CNRS, Paris, 1977).
14. Dastillung, M. & Corbet, M. in *Géochimie Organique des Sédiments Marins Profonds. Orgon II; Atlantique-N. E. Brésil*, 293-323 (CNRS, Paris, 1978).
15. Brassell, S. C. et al. *Init. Rep. DSDP Leg. 56/57*, Part 2, 1367-1390 (1980); in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 375-392 (Pergamon, Oxford, 1980).
16. Wardroper, A. M. K. thesis, Univ. Bristol (1979).
17. Brassell, S. C. thesis, Univ. Bristol (1980).
18. Blumer, M., Mullin, M. M. & Thomas, D. W. *Science* **140**, 974 (1963).
19. Brassell, S. C., Gower, A. P. & Eglinton, G. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 421-426 (Pergamon, Oxford, 1980).
20. Spyckerelle, C., Arpino, P. & Ourisson, G. *Tetrahedron* **28**, 5703-5713 (1972).
21. Albaiges, J., Borbon, J. & Salagre, P. *Tetrahedron Lett.* 595-598 (1978).
22. Han, J. & Calvin, M. *Geochim. cosmochim. Acta* **33**, 739-742 (1969).
23. Haug, P. & Curry, D. J. *Geochim. cosmochim. Acta* **38**, 601-610 (1974).
24. Waples, D. W., Haug, P. & Welte, D. H. *Geochim. cosmochim. Acta* **38**, 381-387 (1974).
25. Kaplan, I. R. & Baedeker, M. J. *Israel J. Chem.* **8**, 529-533 (1970).
26. Cornet, P. A. et al. *Init. Rep. DSDP Leg. 62* (in the press).
27. Whelan, J. K., Hunt, J. M., & Berman, J. *Geochim. cosmochim. Acta* **44**, 1767-1785 (1980).
28. Hammond, D. E., Horowitz, R. M., Broecker, W. S. & Bopp, R. *Init. Rep. DSDP Leg. 20*, 765-771 (1972).
29. Whelan, J. K. & Hunt, J. M. *Init. Rep. DSDP Leg. 56/57*, Part 2, 1349-1365 (1980).
30. Whelan, J. K. & Hunt, J. M. *Init. Rep. DSDP Leg. 50*, 623-624 (1980).
31. Von Huene, R. et al. *Geotimes* **23**(4), 16-21 (1978).
32. Cox, R. E., Maxwell, J. R., Ackman, R. G. & Hooper, S. N. *Can. J. Biochem.* **50**, 1238-1241 (1972).
33. Patience, R. L., Rowland, S. J. & Maxwell, J. R. *Geochim. cosmochim. Acta* **42**, 1871-1875 (1978).
34. Didyk, B. M., Simoneit, B. R. T., Brassell, S. C. & Eglinton, G. *Nature* **272**, 216-222 (1978) and refs therein.
35. Howard, D. L. thesis, Univ. California, Los Angeles (1980).
36. Cardoso, J. N. et al. *Init. Rep. DSDP Leg. 44*, 965-976 (1979).
37. Brassell, S. C. et al. *Init. Rep. DSDP Leg. 50*, 647-664 (1980).
38. Thomson, I. D., Brassell, S. C., Eglinton, G. & Maxwell, J. R. *Init. Rep. DSDP Leg. 64*, (in the press).
39. Curran, J. R. et al. *Geotimes* **24** (7), 18-20 (1979).
40. Simoneit, B. R. T., Mazurek, M. A., Brenner, S., Crisp, P. T. & Kaplan, I. R. *Deep Sea Res.* **26A**, 879-891 (1979).
41. Oremland, R. S., Culbertson, C. & Simoneit, B. R. T. *Init. Rep. DSDP Leg. 64*, (in the press).
42. Belyaev, S. S., Lein, A. Yu. & Ivanov, M. V. in *Biogeochemistry of Ancient and Modern Environments* (eds Trudinger, P. A., Walter, M. R. & Ralph, B. J.) 235-242 (Australian Academy of Science, Canberra, 1980).
43. Romesser, J. A., Wolfe, R. S., Mayer, F., Spiess, E. & Walter-Mauruschat, A. *Arch. Microbiol.* **121**, 147-153 (1979).
44. Farrington, J. W., Frew, N. M., Gswend, P. M. & Tripp, B. W. *Estuarine Coastal mar. Sci.* **5**, 793-808 (1977).
45. Boehm, P. D. & Quinn, J. G. *Estuarine Coastal. mar. Sci.* **6**, 471-494 (1978).
46. Kimble, B. J. thesis, Univ. Bristol (1972).
47. Corbet, B., Albrecht, P. & Ourisson, G. *J. Am. chem. Soc.* **102**, 1171-1173 (1980).

Anomalous ¹³C depletion in early Precambrian graphites from Superior Province, Canada

M. Schoell

Bundesanstalt für Geowissenschaften und Rohstoffe, Stilleweg 2, 3000 Hannover 51, FRG

F.-W. Wellmer*

Metallgesellschaft AG Reuterweg 14, 6000 Frankfurt a.M., FRG

The carbon isotope composition of organic carbon is potentially, at least, a guide to the conditions in which and to the organisms by which it was formed. The isotopic composition of most Phanerozoic kerogens, measured as $\delta^{13}\text{C}$ values, lies within the range -20‰ to -30‰. Most (~65%) samples of Precambrian organic carbon, now represented by graphite and other forms of reduced carbon, have isotopic ratios within the same range, but in others the concentration of ¹³C is significantly depleted by as much as 10‰. The causes of this depletion have an obvious bearing on the conditions of life in the Precambrian. Here we report the results of our measurements of samples of graphite from the Archean of the Superior Province of Canada, one of which being the most depleted in ¹³C ever found in the Precambrian. We explain this ¹³C depletion by organisms in some areas living on CO₂ oxidized from CH₄ that has been formed by methane-producing bacteria in Precambrian stratified seas.

Some 35% of the Precambrian graphite samples measured have ¹³C concentrations which lie outside the Phanerozoic range. Enrichment has usually been explained by the preferential loss of ¹²C during metamorphism¹⁻⁴, but depletion is harder to account for. It has been suggested that depletion may be a consequence of higher CO₂ partial pressure in the Precambrian^{5,6} or may have come about because of the predominance of lipids depleted in ¹³C in Precambrian organic matter⁷.

All ¹³C/¹²C ratios of graphite samples from the Archean of the Superior Province studied so far were within the range of Phanerozoic kerogens⁸. They were from graphites associated with sulphidic and oxidic iron formations and their transition zone. Iron formations are commonly intercalated between volcanoclastic sediments, less common between volcanic flows, and closely associated with cherts, carbonaceous argillites and graywackes. Iron formation horizons can often be followed for tens of kilometres indicating extensive sedimentary basins; the Michipicoten area of northwestern Ontario is a typical example. The isotope ratios of graphites from iron formation horizons of this area all lie in the range -20 to -28‰ (ref. 8). This range also covers the only sample of this report collected in a similar sulphidic iron formation environment of the Savant Lake-Sturgeon Lake area in northwestern Ontario (102 PQ, Table 1).

The other nine samples on which we report were collected from drill holes in a different type of sulphidic graphitic sediments which occur within the volcanic piles themselves. These sediments consist of thin layers of graphitic sulphidic tufts or interflow sediments with strike lengths of the order of only 100-1,000 m. In contrast to the large sedimentary basins with extensive iron formations these short graphitic horizons are imbedded in local depressions the size of which is indicated by the strike lengths of the graphite horizons derived from airborne and ground geophysics. One of these horizons extends over a distance of 1,400 m but all others stretch only a few hundred metres indicating rather small pools.

* Present address: Bundesminister für Wirtschaft Villemombler Straße, 5300 Bonn 1, FRG.

Table 1 Field data and analytical results

Samples	Lat. Long.	Location	S*	$\delta^{13}\text{C}_{\text{PDB}}$ (‰)	Footwall	Geological environment Hangingwall
			L† T‡ (m)			
102 PQ	50°23'N 90°34'W	Conant Township NW Ontario Wabigoon Belt	0 8,000 4	-25.9	Rhyolite tuffs	Intermediate to mafic flows and tuffs
U 6	50°23'N 90°34'W	Conant Township NW Ontario Wabigoon Belt	800 1,400 2.4	-22.5	Dacitic tuffs and shales	Dacitic tuffs and shales
A 3/1§	47°08'N 84°14'W	Township 26, R Range 13 (Olsen) Ontario, Batchawana Wawa Belt	1,200 250 0.9	-26.9	Mafic to intermediate tuffs	Mafic to intermediate tuffs
A 3/2§	47°08'N 84°14'W	Township 26, R Range 13 (Olsen) Ontario, Batchawana Wawa Belt	1,200 250 0.9	-26.9	Mafic to intermediate tuffs	Mafic to intermediate tuffs
S 2	49°16'N 89°42'W	Heaven Lake NW Ontario Wabigoon Belt	800 210 15	-24.6	Cherty sediments	Cherty sediments and massive sulphides
K 1	49°59'N 87°42'W	Onaman Lake area NW Ontario Wabigoon Belt	400 270 3.1	-21.0	Mafic fragmentals	Siltstone and mafic tuffs
K 10	49°59'N 87°42'W	Onaman Lake area NW Ontario Wabigoon Belt	400 270 1.7	-25.2	Mafic fragmentals	Siltstone and mafic tuffs
K 2¶	49°59'N 87°42'W	Onaman Lake area NW Ontario Wabigoon Belt	1,100 150 4.7	-35.5	Tuffaceous shales	Tuffaceous shales
K 6	49°50'N 78°41'W	Sandra Township NW Ontario Wabigoon Belt	900 275 3.3	-38.3	Rhyodacite	Andesite
X-1	48°50'N 78°41'W	Languedoc Township Quebec, Abitibi Belt	700 360 5.0	-43.8	Interbedded mafic to intermediate tuffs and shales	Mafic flows

The samples were crushed and treated with diluted acid to remove possible carbonate material, and then burned at temperatures of 900°C in a stream of pure dried oxygen using a copper oxide catalyst. The resulting CO₂ was analysed for its carbon isotopic composition with a triple collecting mass-spectrometer (VG 903). Usual corrections were applied and the isotopic composition is reported as δ -value versus the PDB-standard. All samples were processed twice. The reproducibility as determined from double measurements is $\pm 0.2\%$. The NBS graphite standard has been determined at $-28.2 \pm 0.1\%$. (δ -values are defined as relative deviations from the PDB standard where $\delta^{13}\text{C} = (R_{\text{sample}}/R_{\text{standard}} - 1) \times 1,000\%$ where R is the isotope ratio $^{13}\text{C}/^{12}\text{C}$.)

* Stratigraphic distance to top of volcanic pile.

† Length of graphite horizons.

‡ Thickness of graphite horizons.

§ Two graphite horizons in the same drill hole separated by a 15 m stratigraphic section.

|| K 1 and K 10 same graphite horizon 60 m apart along strike.

¶ Horizon of K 2, 700 m stratigraphically lower than K 1 and K 10.

The volcanic and sedimentary host rocks of the graphite horizons have been metamorphosed in greenschist facies conditions. Typical mineral assemblages in the volcanics are chlorite, actinolite, epidote, albite, quartz and sericite. Samples A3/1, A3/2 and X-1 are from the Abitibi-Wawa volcanic plutonic belt for which radiometric data indicate an age of $\sim 2,750$ Myr; all other samples are from the 2,800-Myr old Wabigoon volcanic plutonic belt^{9,10}. Our graphite samples come from bands and layers intricately intercalated with volcanoclastic sediments, indicative of their indigenous nature. Microscopic examination shows that graphites form flakes which are commonly oriented parallel to the schistosity; optically the flakes display a clearly visible pleochroism which proves the graphitic nature of the carbonaceous matter.

Table 1 gives the results of our isotope analyses and some field data, obtained from exploration drilling. The geological position of the graphite horizons is estimated from their stratigraphic distance to the top of the volcanic pile which varies between 1,200 and 0 m, mostly being 400–900 m.

Seven of the 10 samples analysed revealed $\delta^{13}\text{C}$ values between -21 and -27% —the normal range within the spread of data of many other Precambrian samples (Fig. 1). Three samples show significant depletion in the heavy isotope with $\delta^{13}\text{C}$ values between -35% and as low as -43.8% ; this value is among the lowest ^{13}C concentration ever reported for Precambrian reduced carbon. The field data indicate a common geological setting even for those localities which yielded the anomalously low ^{13}C graphite data. Sample K2 from the Onaman Lake area in northwestern Ontario with an extremely low $\delta^{13}\text{C}$ value was taken in the same volcanic pile only 700 m away from two normal value samples (K1 and K10).

Anomalous ^{13}C depletion ($\delta^{13}\text{C}$ less than -30%) in Precambrian organic matter has been reported previously^{4,5,12,14,16} (Fig. 1). But some of the values shown in Fig. 1 are not indicative of the primary source. For example, $\delta^{13}\text{C}$ values of -37.01 and -39.1% respectively have been reported for the so-called anthraxolites⁴. Stringers of visible carbon in limestones were also found to be depleted in ^{13}C ($\delta^{13}\text{C} \approx -37$ to -41%) as

compared with disseminated carbon which had 'normal' values around -26 to -33% (ref. 16). Secondary effects must have been operative to form both the anthraxolites and the stringers which forbids any conclusion as to their original isotopic composition. Such secondary effects are ruled out for our graphite samples (data set 1 in Fig. 1). They are undoubtedly of syngenetic nature which is indicated by the above described primary sedimentary features.

Thus the anomalous $^{13}\text{C}/^{12}\text{C}$ isotope ratios indicate special environmental conditions which must have prevailed in some of these small depressions. These special conditions are assumed to have prompted the organic species to incorporate the ^{12}C -isotope more readily.

We discarded the theory that in some pools specific assemblages of organisms had different isotope fractionations when assimilating carbon. Also, we consider it unlikely that various differing metabolic systems were involved.

The change of the isotopic composition of the carbon source is therefore considered in the following as the process which may easily explain such a large change in organic $^{13}\text{C}/^{12}\text{C}$ ratios as we observed or even larger changes. In all cases the formation of graphite is closely interrelated with the occurrence of sulphidic sediments indicating areas of hydrothermal activity. The discharge of hydrothermal brines into morphological depressions is likely to cause a stratification of waters as described for the Red Sea brine pools, the Lake Kivu or more recently from below the Gulf of Mexico¹⁷⁻²⁰. The brines prevent vertical mixing and thereby act as traps for organic debris. Methane can form in such environments and may have become trapped in the stratified waters. The formation of methane by bacteria is the only geochemical process known to bring about appreciable ^{13}C -depletion (typical $\delta^{13}\text{C}$ values are around -60 to -80%). Methane may become oxidized to CO_2 when appreciable residence times within the stratified waters allow CH_4 oxidizing organisms to grow²⁴. The CO_2 from methane oxidation is even more depleted in ^{13}C than the methane²¹ and its ^{13}C concentration can range from -70 to -90% . Such ^{13}C -depleted CO_2 would then be admixed to CO_2 which may have been derived

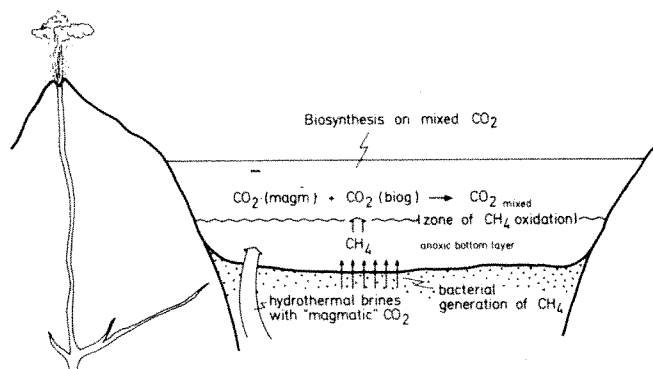


Fig. 2 Hypothetical explanation of the ^{13}C depletion in Precambrian organic carbon. Stratified waters are traps for methane which is bacterially produced in the stratified waters. Oxidation in the sediment or at water interfaces would generate ^{13}C depleted CO_2 as a carbon source for metabolizing eukaryotic biota.

from hydrothermal solutions or from air CO_2 to form a carbon reservoir, depleted in ^{13}C (Fig. 2). Oxidation of CH_4 to CO_2 does not necessarily need free atmospheric oxygen because anaerobic methane oxidation is performed in nature by sulphate reducers^{22,23}:



and sulphate ions are likely to have been available from seawater or hydrothermal brines.

Methane-producing bacteria are ubiquitous in nature and must be regarded as one of the earliest species of organized life on Earth. They may, therefore also have played an important role in Precambrian times. The above model, summarized in Fig. 2, attempts to explain the as yet not understood variability of $^{13}\text{C}/^{12}\text{C}$ ratios in Precambrian reduced carbon. The extent of ^{13}C depletion would depend on the mixing ratio of normal CO_2 and CO_2 oxidized from CH_4 . An addition of 20% oxidized CO_2 to a CO_2 pool with an initial $\delta^{13}\text{C}$ of -7% would change the organic carbon $\delta^{13}\text{C}$ by 14% in favour of the light isotope, that is, the isotopic composition of -43% found in one graphite sample would have required an addition of $\sim 25\%$ of CO_2 which has been produced from CH_4 oxidation. Whether or not CH_4 and the CO_2 oxidized from it enters the CO_2 pool in which organic matter is metabolized depends on the specific situation of each basin. In shallow waters without stratification methane escapes without being oxidized; in well stratified waters methane is oxidized during diffusional exchange at density interfaces or alternatively at the sediment-water interface²⁰. Note that a relatively large fraction of our graphite samples (short non-iron formation horizons within the volcanic pile) is depleted in ^{13}C (three out of nine). In the graphites from large basins, no ^{13}C -depleted samples were found⁸. This may be due to differences in basin dimensions: in small basins stratified waters are more stable than in large basins.

The above processes are not necessarily confined to the small basins of the Archean of the Superior Province. They could have operated in other Precambrian areas where anomalous ^{13}C -depleted reduced carbon is found. The bulk of reduced Precambrian carbon has the same ^{13}C concentration as Phanerozoic kerogens and the ^{13}C -depleted reduced carbon might be explained by local effects. This reasoning argues against the interpretation of low ^{13}C concentrations in Precambrian organic matter as being indicative of special conditions in the Precambrian as compared with Phanerozoic times. The completely different environmental conditions (low O_2 partial pressure) obviously had not affected the basic processes of CO_2 assimilation.

We thank Teck Explorations Ltd for permission to publish the data.

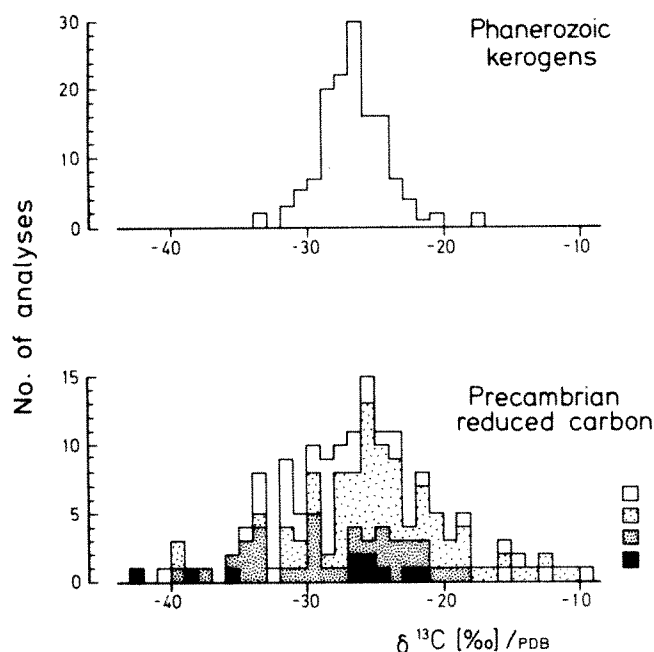


Fig. 1 Frequency distribution of $\delta^{13}\text{C}$ values of Precambrian reduced carbon compared with Phanerozoic kerogens. 1, This paper; 2, graphites from various Precambrian iron-formations^{4,5,8,11,12,14,15}; 3, disseminated carbon and carbon stringers from the middle Precambrian Transvaal dolomite¹⁶; 4, various Precambrian formations¹³. Phanerozoic kerogen values are unpublished data of the BGR stable isotope laboratory.

Received 1 September 1980; accepted 26 January 1981.

- Hoering, T. C. in *Researches in Geochemistry* (ed. Abelson, P. H.) (Wiley, New York, 1967).
- Rumble, D., Hoering, T. C. & Grew, E. S. *Yb. Carnegie Instn Wash.* **76**, 623 (1977).
- McKirdy, D. M. & Powell, T. G. *Geology* **2**, 591 (1974).
- Barghoorn, E. S., Knoll, A. H., Dembicki, H. & Meinschein, W. G. *Geochim. cosmochim. Acta* **41**, 425 (1977).
- Smith, J. W., Schopf, J. W. & Kaplan, I. R. *Geochim. cosmochim. Acta* **34**, 659 (1970).
- Degens, E. T. in *Organic Geochemistry* (eds Eglinton, G. & Murphy, M. T. J.) 304 (1969).
- Galimov, E. in *Kerogen Insoluble Organic Matter from Sedimentary Rocks* (ed. Durand, B.) 271 (Edition Technique, Paris, 1980).
- Goodwin, A. M., Monster, J. & Thode, H. G. *Econ. Geol.* **71**, 870 (1976).
- Krogh, T. E. & Davis, G. L. *Yb. Carnegie Instn Wash.* **70**, 241 (1971).
- Nunes, P. D. & Thurston, P. D. *Can. J. Earth Sci.* **17**, 710 (1980).
- Hoering, T. C. *Yb. Carnegie Instn Wash.* **61**, 190 (1962).
- Cloud, P. E. Jr, Gruner, J. W. & Hagen, H. *Science* **148**, 1713 (1965).
- Oehler, D. Z., Schopf, J. W. & Kvendovden, K. A. *Science* **175**, 1246 (1972).
- Perry, E. C. Jr, Tan, F. C. & Morey, G. B. *Econ. Geol.* **68**, 1110 (1973).
- Schidlowski, M., Appel, P. W. U., Eichmann, R. & Junge, C. E. *Geochim. cosmochim. Acta* **43**, 189 (1979).
- Eichmann, R. & Schidlowski, M. *Geochim. cosmochim. Acta* **39**, 585 (1975).
- Schoell, M. & Hartmann, M. *Mar. Geol.* **14**, 1 (1973).
- Tietze, K., Geyh, M., Müller, H., Schröder, L. & Stahl, W. *Geol. Rdsch.* **69**, 452–472 (1980).
- Brooks, T. M., Bright, T. J., Bernard, B. B. & Schwab, C. R. *Limnol. Oceanogr.* **24**, 735 (1979).
- Sackett, W. M. *et al. Earth planet. Sci. Lett.* **44**, 73 (1979).
- Coleman, D. D., Risatti, J. & Schoell, M. (in preparation).
- Barnes, R. O. & Goldberg, E. D. *Geology* **4**, 297 (1976).
- Martens, C. S. & Berner, R. A. *Limnol. Oceanogr.* **22**, 10 (1977).
- Fallon, R. D., Havrilt, S., Hanson, R. S. & Brock, T. D. *Limnol. Oceanogr.* **25**, 357 (1980).

Population density and body size in mammals

John Damuth

Committee on Evolutionary Biology, University of Chicago,
1103 E. 57 Street, Chicago, Illinois 60637, USA

There seems to be an inverse relationship between the size of an animal species and its local abundance. Here I describe the interspecific scaling of population density and body mass among mammalian primary consumers (herbivores, broadly defined). Density is related approximately reciprocally to individual metabolic requirements, indicating that the energy used by the local population of a species in the community is independent of its body size. I suggest that this is a more general rule of community structure.

Figure 1 shows the logarithm of mean population density plotted against the logarithm of mean adult body mass for 307 species of mammalian primary consumers. As far as possible, the density values represent 'ecological' densities, that is, those which apply to the habitat area actually used by the species. The

relationship is linear, with a slope of -0.75 , and species densities seem to be restricted to varying within about one order of magnitude from the value predicted by the regression line.

This analysis combines data from a wide variety of habitats throughout the world, and it is important that we know whether this overall pattern accurately reflects that found within individual communities; few, if any, mammal communities have been completely studied. To obtain a representative sample, I extracted from my data those sets yielding densities of three or more species within the same habitat type; these 'constructed' communities are shown in Table 1, with regression statistics. They include representatives of almost the whole range of habitat structures and primary productivity levels encountered by terrestrial mammals. None of the individual slopes, which vary from -0.56 to -0.95 , differ significantly from -0.75 , and Levene's test¹ and an analysis of covariance² reveal no significant differences in variance about the regression, in slope, or in intercept. Pooling the data from the communities gives an estimated slope of -0.70 , which is not significantly different from the value obtained for the overall regression (t -test). Thus, the overall trend gives a reasonable representation of that which we are likely to find within individual communities.

Knowledge of population density scaling in communities allows us to consider the relative energy use of species among primary consumers of different sizes. Individual basal metabolic requirements are related to body mass by the power of 0.75 (refs 3, 4). Estimates of the metabolic requirements of free-living mammals in natural habitats roughly parallel basal requirements, but at a higher level, varying between ~ 1.5 and 3.0 times basal values^{5–13}. Thus they will also be related to body mass by the power of ~ 0.75 . The energy used by the local population of a species equals the population density (D) multiplied by individual metabolic requirements (R), which yields the following relationship to body mass (W): $DR \propto W^{-0.75} W^{0.75}$. The exponents of W cancel each other, which gives the important result that the amount of energy that a species population uses in the community is independent of its body size. No mammal herbivore species, on an ecological time scale, has an energetic advantage over any other solely as a result of size differences.

There are two important corollaries of this relationship. First, because secondary productivity for a given amount of assimilated energy is independent of body size in mammals (Kleiber's Law)³, it follows from the above that the secondary productivity of a herbivore species' local population, and hence the energy that it yields to the next higher trophic level, is also independent of body size. Second, the standing-crop biomass of a species (population density multiplied by body mass) is positively related to body mass by the ~ 0.25 power; this is a

Table 1 Habitat types and constructed communities

Community/habitat type	a	s.e.	r	b	n
Sonoran desert, USA	-0.63	0.16	-0.76	3.37	14
Mesquite grassland, USA and Mexico	-0.56	0.086	-0.95	3.73	7
Boreal and subalpine forest, N. America	-0.79	0.080	-0.95	4.43	12
Lowland tropical rainforest, Malaysia	-0.60	0.089	-0.90	2.61	13
Transvaal lowveld (woodland-savanna), S. Africa	-0.61	0.089	-0.93	3.78	10
Mixed temperate forest, Poland	-0.79	0.080	-0.97	4.33	8
Temperate grassland, USA	-0.67	0.13	-0.92	3.28	7
Tropical grassland, Rwenzori N.P., Uganda	-0.79	0.17	-0.81	3.59	13
Tropical grassland, Sri Lanka*	-0.79	0.19	-0.92	4.59	5
Ichu grassland, Altiplano, Peru*	-0.82	0.34	-0.99	3.93	4
High arctic tundra, Canada*	-0.95	0.21	-0.96	4.71	4
Sub-arctic birch forest and meadows, Norway*	-0.83	0.72	-0.99	4.41	3
Southern pine-hardwood forest, USA*	-0.91	0.46	-0.89	4.23	3
Northern hardwoods, USA*	-0.57	0.11	-0.98	3.11	3
Oaks and chaparral, USA*	-0.72	0.021	-0.99	4.83	3
Means of statistics	-0.74	0.20	-0.93	3.93	—
Pooled data from above regressions	-0.70	0.040	-0.86	3.71	109

Statistics are for standard least-squares regression equations: $\log D = a(\log W) + b$; s.e., standard error of the slope (a); r, correlation coefficient; n, number of species. Note that in the pooled regression, as some species are found in more than one habitat type, $n = 109$ includes only 92 separate species.

* Due to small sample size and/or a single point at one end of the size range, the individual values for these regressions are not very reliable.

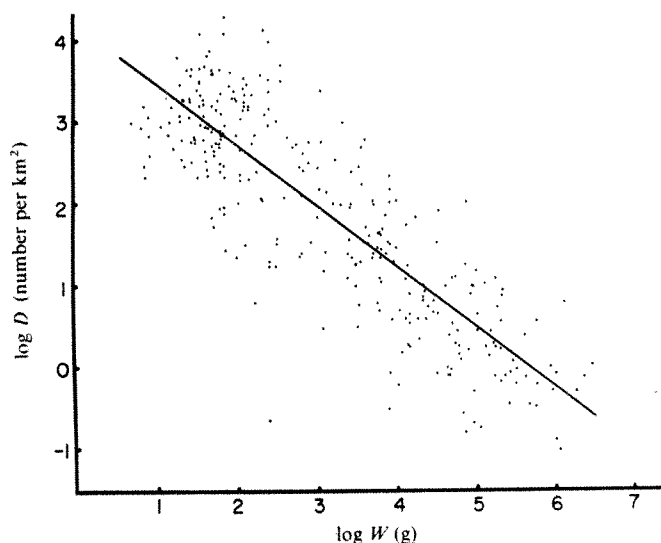


Fig. 1 Population density (D) compared with the mean adult body mass (W) for 307 mammal primary consumers; each point represents one species. Density values for each species are the mean of the means from each locality from which data were reported for the particular species. (Data are from the literature for the years 1950–79, derived from 115 journals and numerous books, ~650 references in all.) The line represents the least-squares regression line, $\log D = -0.75 (\log W) + 4.23$; $r = -0.86$, standard error of the slope = 0.026.

quantitative measure of a qualitative relationship that has been known for some time but only sporadically appreciated¹⁴. Species of small mammals are able to produce, on average, the same amount of biomass over time as do species of large mammals, whereas at a given moment their standing-crop biomass is considerably less, because the population turnover rates and individual growth rates per unit weight of small species are much greater^{15–17}.

The widespread occurrence of an approximately reciprocal relationship between population density and individual metabolic requirements among mammalian herbivores suggests that a general principle is involved. The independence of species energy control and body size revealed by this reciprocal relationship implies that random environmental fluctuations and interspecific competition act over evolutionary time to keep energy control of all species within similar bounds. It is unlikely that the occurrence of competition among species of different body sizes is restricted to herbivorous mammals. Along with the ubiquity of allometric scaling of metabolic rate with body size⁴, this suggests that a reciprocal relationship, and particularly a value of ~ -0.75 , characterizes a broader range of taxa and trophic levels. These points will be discussed more fully elsewhere.

I thank I. L. Heisler, V. C. Maiorana, D. M. Raup, V. L. Roth and L. Van Valen for comments, and acknowledge support from the University of Chicago and Field Museum of Natural History.

Received 6 October 1980; accepted 27 January 1981.

1. Van Valen, L. *Evol. Theory* **4**, 33–44 (1978).
2. Brownlee, K. A. *Statistical Theory and Methodology in Science and Engineering* 2nd edn (Wiley, New York, 1965).
3. Kleiber, M. *The Fire of Life* 2nd edn (Krieger, New York, 1975).
4. Hemmingen, A. M. *Rep. Steno meml Hosp.* **4**, 7–58 (1950).
5. Brody, S. *Bioenergetics and Growth* (Reinhold, New York, 1945).
6. Chew, R. M. & Chew, A. E. *Ecol. Monogr.* **40**, 1–21 (1970).
7. Gessaman, J. A. in *Ecological Energetics of Homeotherms* (ed. Gessaman, J. A.) (Utah University Press, 1973).
8. King, J. R. in *Avian Energetics* (ed. Paynter, R. A. Jr) (Nuttall Ornithological Club, Massachusetts, 1974).
9. McKay, G. M. *Smithson. Contr. Zool.* **125**, 1–113 (1973).
10. Moen, A. N. *Wildlife Ecology: an Analytical Approach* (Freeman, San Francisco, 1973).
11. Mullen, R. K. *Comp. Biochem. Physiol.* **39**, 379–390 (1971).
12. Mullen, R. K. & Chew, R. M. *Ecology* **54**, 633–637 (1973).
13. Nagy, K. A. & Milton, K. *Ecology* **60**, 475–480 (1979).
14. Odum, E. P. *Fundamentals of Ecology* 3rd edn (Saunders, Philadelphia, 1971).
15. Fenichel, T. *Oecologia* **14**, 317–326 (1974).
16. Western, D. *Afr. J. Ecol.* **17**, 185–204 (1978).
17. Case, T. J. *Q. Rev. Biol.* **53**, 243–282 (1978).

Competitive ability influences habitat choice in marine invertebrates

Richard K. Grosberg

Department of Biology, Yale University, 260 Whitney Avenue, PO Box 6666, New Haven, Connecticut 06511, USA

Patterns of distribution and abundance of sessile marine epibenthic invertebrates are controlled by three factors: (1) the presence and abundance of larvae which are competent to settle, (2) the choice of settling sites by recruiting larvae, and (3) the biotic and physical events occurring during and after settlement. Although there is much information on the distribution of larvae and seasons of recruitment^{1–3}, substratum selection^{4–6} and post-settlement events^{7–15}, very little is known of the ecological and evolutionary relationships between these factors^{5,16–18}. Natural selection acts on entire life cycles, thus information about these relationships is essential for understanding patterns of recruitment and survival. For example, sessile organisms can modify the course of post-recruitment events by selective settlement and directional growth¹⁹. Here, I present evidence that the larvae of several taxa of marine invertebrates avoid substrata where there is a high probability of death caused by a superior spatial competitor.

I found, as did Grave²⁰, that in the Eel Pond at Woods Hole, Massachusetts (USA), the two compound ascidians *Botryllus schlosseri* (Pallas) Savigny and *Botrylloides leachii* (Savigny) cover, by area, over 50%, and sometimes 100%, of hard substrata during the spring and summer. This period includes the season of recruitment for most other colonizers. Thus, settling larvae are likely to contact a botryllid ascidian at some time during their lives. A common result of contacts between colonists is overgrowth of one organism by another^{21–29}. In this way, overgrowth is a consequence of competition for space (and perhaps other resources³⁰), and therefore can be an important factor in the distribution and abundance of encrusting species.

B. schlosseri is the most successful overgrower (along with *B. leachii*) in the Eel Pond (Table 1) and is rarely overgrown by other species. As it is such an important member of the Eel Pond epibenthic assemblage (considering the post-settlement events of overgrowth and coverage), I investigated the effects of its density on the recruitment of other taxa.

B. schlosseri collected from the Eel Pond were placed in finger-bowls where they released larvae which were then transferred to 10-ml polycarbonate dishes. Two hours after settlement on the sides of the dishes, the juveniles were teased off and re-attached to 5 × 5 cm glass plates. I set up three densities of juvenile *Botryllus* with three replicates of each: (1) 15 regularly spaced *Botryllus* juveniles, with all subsequent botryllid settlement removed daily; (2) 5 regularly spaced *Botryllus* juveniles with all subsequent botryllid recruits removed daily; and (3) no *Botryllus* juveniles with daily removal of all botryllid recruits. Three additional glass plates which initially carried no juveniles, but on which all recruits were allowed to remain, served as controls.

Experimental and control plates were positioned horizontally in a Latin square arrangement³¹ below a floating dock at 1 m depth. The experiment was started on 20 July, 24 h after the larvae had been transplanted. The lower side of each plate (which carried the juvenile *Botryllus*) was censused non-destructively until 28 July on a daily basis—this allowed me to distinguish between failure to settle and all but the earliest post-settlement mortality.

Figure 1 shows the numbers of each species that settled on the glass plates. To determine if there were any differences between each of the three density treatments and the controls, the data for each species were analysed using the null hypothesis that

settlement was uniform in each case. χ^2 tests unambiguously placed the species in two groups: (1) selective species or those for which settlement was not uniform between treatments (Fig. 1a, all $P < 0.005$) and (2) non-selective species or those for which there were no significant differences in settlement between density treatments (all $P > 0.2$). When settlement on the plates carrying five *B. schlosseri* was compared with that on plates where all botryllids had been removed, settlement was uniform for all species ($P > 0.5$). On control plates, where cumulative *B. schlosseri* settlement reached 15 per plate on the third day of the experiment, settlement of selective species was intermediate between that on plates carrying 15 *B. schlosseri* and those carrying none or five juveniles. This implies that the threshold for detection of *B. schlosseri* by selective species lies somewhere between 5 and 15 resident *Botryllus* juveniles.

There is a clear relationship between patterns of recruitment and a species' susceptibility to overgrowth. The species which did not settle on plates where *Botryllus* number was 15 (selective

species) were also overgrown by *Botryllus* in 1,008 of 1,048 contacts (96.2%). These species are the first nine of Table 1. In contrast, those species which settled uniformly, regardless of *B. schlosseri* density (non-selective species), were rarely overgrown by *Botryllus* (171 of 961 contacts or 17.1%). These species include the remaining 10 of Table 1.

The mechanisms by which settlement is decreased for some taxa, but not for others, are unclear. Allelopathy^{22,32,33} cannot account for these results, for when selective species do settle on the plates (for example, at low *Botryllus* densities), they often settle immediately adjacent to the competitive dominant. In addition, because *Botryllus* cannot ingest particles as large as most recruiting larvae^{34,35}, selective predation does not explain the results. It is more likely that the settling larvae of some taxa can recognise the density of resident *Botryllus*, and reject substrata when that density exceeds a certain threshold.

Selective settlement has been observed in a number of taxa, but it is not understood which factors make particular substrata

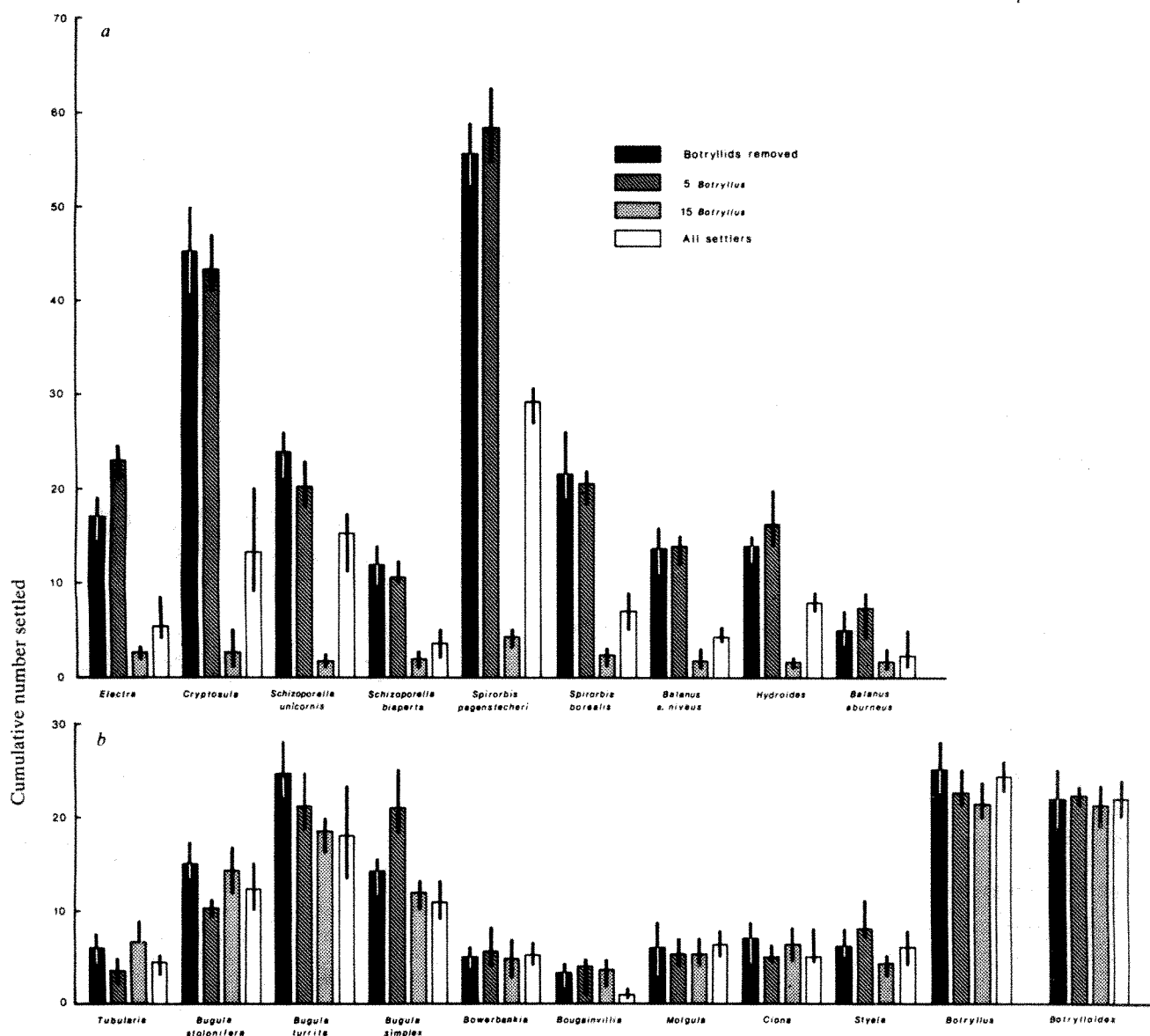


Fig. 1 The cumulative settlement for all species which settled during the experiment. Bars represent the mean settlement for the three replicates of each density treatment and the controls. Lines indicate the range of settlement values. The plates were arranged in a Latin square so that plate position effects could be detected: the χ^2 analyses indicated that there were no significant differences between replicates, thus there were no plate position effects. Of the 60 transplanted larvae, only one died during the experiment. Any settler not immediately identified was allowed to grow after the term of the experiment until identification was possible. At the end of the experiment, none of the colonists covered $>2 \text{ mm}^2$, thus coverage was low. *a*, Data for species which showed a settlement effect at different *B. schlosseri* densities. These are termed selective species—this group includes four encrusting bryozoans, three tube-building polychaetes and two barnacles. All these species have feeding structures close to the substratum, thus easily obstructed by overgrowth. *b*, Data for those species which showed no significant differences in settlement at different *Botryllus* densities. These are termed non-selective species—this group includes four arborescent bryozoans, two stoloniferous hydroids, three solitary ascidians and two colonial botryllid ascidians. With the exception of the botryllids, all members of this group have elevated feeding structures and therefore are generally resistant to the effects of overgrowth (see refs 23, 28).

Table 1 Results of contacts between *Botryllus schlosseri* and the adults of the species which recruited during the period 20–28 July 1979

Overgrown	Wins	Ties	Total	Wins/total (%)	Losses/total (%)
<i>Electra crustulenta</i>	76	0	78	76/78 (97.4%)	2/78 (2.6%)
<i>Cryptosula pallasiana</i>	159	0	159	159/159 (100%)	0/159 (0%)
<i>Schizoporella biaperta</i>	43	0	43	43/43 (100%)	0/43 (0%)
<i>Schizoporella unicornis</i>	58	0	63	58/63 (92.1%)	5/63 (7.9%)
<i>Spirorbis pagenstecheri</i>	382	0	382	382/382 (100%)	0/382 (0%)
<i>Spirorbis borealis</i>	118	13	131	118/131 (90.1%)	0/131 (0%)
<i>Balanus amphitrite niveus</i>	79	10	89	79/89 (88.8%)	0/89 (0%)
<i>Balanus eburneus</i>	47	5	52	47/52 (90.4%)	0/52 (0%)
<i>Hydroides dianthus</i>	46	5	51	46/51 (90.2%)	0/51 (0%)
<i>Botrylloides diegensis</i>	147	108	391	147/391 (37.6%)	136/391 (34.8%)
<i>Tubularia larynx</i>	3	87	84	3/87 (3.5%)	0/87 (0%)
<i>Bougainvillia</i> sp.	2	32	34	2/34 (5.9%)	0/34 (0%)
<i>Bugula stolonifera</i>	2	59	61	2/61 (3.3%)	0/61 (0%)
<i>Bugula turrita</i>	4	76	80	4/80 (5.0%)	0/80 (0%)
<i>Bugula simplex</i>	3	99	102	3/102 (2.9%)	0/102 (0%)
<i>Bowerbankia gracilis</i>	6	64	70	6/70 (8.6%)	0/70 (0%)
<i>Molgula manhattensis</i>	1	26	27	1/27 (3.7%)	0/27 (0%)
<i>Ciona intestinalis</i>	1	15	16	1/16 (6.3%)	0/16 (0%)
<i>Styela clava</i>	2	95	97	2/97 (2.1%)	0/97 (0%)

These data were collected during the summer of 1979 from contacts between animals living on algae, floating docks, rocks and glass plates in the Eel Pond. All contacts were followed from their initiation until their resolution by overgrowth or mutual retreat. Total contacts equal wins + losses + ties. A win is scored when *Botryllus* overtops and kills >50% of a colonial organism, or entirely kills a solitary organism. A tie is scored when growth of both participants ceases along a contact, or when *Botryllus* simply grows around another colonizer without perceptible harm to either. Although *Botryllus* infrequently overgrows certain taxa, it is rarely overgrown itself by any taxa. Nevertheless, the canopies of arborescent bryozoans and hydroids sometimes shade underlying compound ascidians.

(or habitats) better than others. Notable exceptions include: (1) the gregarious settlement of some barnacles which apparently ensures cross-fertilization and viability of eggs^{6,33,36}; (2) the preferential settlement of predators, commensals and parasites near their prey, partners and hosts, respectively⁵; (3) the selection by some bryozoans and polychaetes of particular algae (or positions on these algae) which correlates with patterns of survival of juveniles and adults^{17,18,37}; and (4) the gregarious settlement of some arborescent bryozoans which enhances their interspecific competitive ability³⁸.

The results from the Eel Pond indicate that some larvae can distinguish substrata where post-settlement events are likely to kill them. Apparently, these larvae avoid encounters with the competitive dominant, *B. schlosseri*, by settling away from dense stands of it. In comparison, the larvae that do settle where *B. schlosseri* is abundant are those most immune to overgrowth by it. Taken together, the data illustrate that competition during the post-settlement phase of an organism's life cycle can strongly influence the evolution of habitat selection.

This research was supported by the Lerner Fund for Marine Research of the American Museum of Natural History. The Woods Hole Oceanographic Institution provided laboratory space. I thank L. Buss, R. Harrison, R. Hartigan, B. Keller, B. Kennedy and N. Knowlton for their helpful criticisms.

Received 29 October 1980; accepted 28 January 1981.

- MacDougall, K. D. *Ecol. Monogr.* **13**, 321–374 (1943).
- Thorson, G. *Neth. J. Sea Res.* **3**, 267–293 (1966).
- Osman, R. W. *Ecol. Monogr.* **47**, 37–63 (1977).
- Meadows, P. S. & Campbell, J. I. *Adv. mar. Biol.* **10**, 271–382 (1972).
- Schelteima, R. S. *Thalassia jugosl.* **10**, 263–296 (1974).
- Crisp, D. J. in *Adaptations to Environment: Essays on the Physiology of Marine Organisms* (ed. Newell, R. C.) 83–124 (Butterworths, London, 1976).
- Connell, J. H. *Ecology* **42**, 710–723 (1961).
- Connell, J. H. *Ecol. Monogr.* **31**, 61–104 (1961).
- Paine, R. T. *Ecology* **50**, 950–961 (1969).
- Paine, R. T. *Oecologia* **15**, 93–120 (1974).
- Dayton, P. K. *Ecol. Monogr.* **41**, 315–389 (1971).
- Connell, J. H. & Slatyer, R. O. *Am. Nat.* **111**, 1119–1144 (1977).
- Menge, B. A. *Ecol. Monogr.* **46**, 355–393 (1976).
- Sutherland, J. P. & Karlson, R. H. *Ecol. Monogr.* **47**, 425–446 (1977).
- Buss, L. W. & Jackson, B. C. *Am. Nat.* **113**, 223–234 (1979).
- Woodin, S. A. *J. mar. Res.* **34**, 25–41 (1976).
- Mackay, T. F. C. & Doyle, R. W. *Heredity* **40**, 1–12 (1978).
- Harward, P. J. & Harvey, P. H. *J. mar. biol. Ass. U.K.* **54**, 677–684 (1974).
- Buss, L. W. in *Biology and Systematics of Colonial Organisms* (ed. Larwood, G. P.) 459–497 (Academic, London, 1979).
- Grave, B. H. *Biol. Bull.* **65**, 375–386 (1933).
- Gordon, D. P. *J. nat. Hist.* **6**, 503–514 (1972).
- Bryan, P. G. *Micronesica* **9**, 237–242 (1973).
- Stebbing, A. R. D. *J. mar. biol. Ass. U.K.* **53**, 247–261 (1973).
- Stebbing, A. R. D. in *Living and Fossil Bryozoa* (ed. Larwood, G. P.) 173–183 (Academic, London, 1973).

- Jackson, J. B. C. *Am. Nat.* **111**, 743–767 (1977).
- Jackson, J. B. C. *J. Anim. Ecol.* **48**, 805–823 (1979).
- Karlson, R. H. *J. exp. mar. Biol. Ecol.* **31**, 225–239 (1978).
- Russ, G. R. *J. exp. mar. Biol. Ecol.* **42**, 55–69 (1980).
- Sebens, K. P. *Ecology* (in the press).
- Buss, L. W. *Nature* **281**, 475–477 (1980).
- Sokal, R. R. & Rohlf, F. J. *Biometry* (Freeman, San Francisco, 1969).
- Goodbody, I. *Nature* **190**, 282–283 (1961).
- Knight-Jones, E. W. & Moyse, J. *Symp. Soc. exp. Biol.* **15**, 72–95 (1961).
- Lorenzoni, G. G. *Atti Ist. veneto* **126**, 147–168 (1968).
- Milkman, R. *Biol. Bull.* **132**, 229–243 (1967).
- Barnes, H. & Crisp, D. J. *J. mar. biol. Ass. U.K.* **35**, 631–639 (1956).
- Knight-Jones, E. W., Knight-Jones, P. & Al-Ogily, S. M. in *9th Eur. mar. biol. Symp.* (ed. Barnes, H.) 539–561 (Aberdeen University Press, 1975).
- Buss, L. (in review).

Fertilization potential in golden hamster eggs consists of recurring hyperpolarizations

Shun-ichi Miyazaki & Yukio Igusa

Department of Physiology, Jichi Medical School, Tochigi-ken 329-04, Japan

The fertilization potential, or activation potential, has been demonstrated in the eggs of various species¹, and it has been shown to block polyspermy in echinoderm^{2,3}, echinuran⁴ and frog⁵ eggs, but no studies have been reported of electrical phenomena occurring when mammalian eggs are fertilized. We report here the fertilization potential of golden hamster eggs *in vitro*. To correlate the change of potential with the interaction between sperm and egg, only one sperm was attached to each egg. We found that a sperm induces recurring hyperpolarizations, constituting a fertilization potential which differs from that in the eggs of other species.

Mature eggs were collected from the oviducts of superovulated females⁶ and freed from the surrounding cumulus cells and zona pellucida by sequential treatment with 0.1% hyaluronidase (~5 min at 23–26°C) and 0.1% trypsin (4–5 min) in medium used for rat eggs⁷. Removal of the zona allows sperm to reach the plasma membrane soon after insemination. Spermatozoa obtained from the cauda epididymis were activated with acrosome reaction by incubation in modified Tyrode solution containing 20% human serum⁸ and 1% bovine adrenal gland extract⁹ at 37°C for 4 h. All treatments and experiments were performed in a 0.4 ml drop of medium placed in a plastic Petri dish and covered with paraffin oil. A glass microelectrode was

inserted in an egg viewed in a phase-contrast inverted microscope and a small amount of the sperm suspension was added to the egg medium at 30–32 °C (final concentration 2,000–5,000 sperm per 0.4 ml). Electrode resistances were in the range 40 to 80 M Ω , and leakage current of a preamplifier was <2 pA. Each experimental egg was transferred quickly to sperm-free medium after withdrawing the electrode, and was incubated for 4 h to determine sperm entry by observation of male pronuclei.

The resting potential of unfertilized eggs averaged -29 ± 7 mV (mean $\pm$ s.d., $n = 120$), but 7% of the values were between -40 and -50 mV. The current–voltage relationship (I – V curve) was linear for potentials of between -100 and 0 mV, and the effective membrane resistance averaged 142 ± 65 M Ω ($n = 98$). The more negative resting potential tended to be associated with the larger membrane resistance, suggesting that there was a substantial leakage due to electrode impalement in most cases. The largest resting potential was -50 mV with a membrane resistance of 400 M Ω , which is estimated to be 66 k Ω cm ^{2} , assuming that the egg is a sphere of 72 μ m diameter.

Figure 1A shows the fertilization potential in the case of the severe polyspermy which usually occurs in zona-free eggs. The major potential change is a repetitive, transient hyperpolarization (H-response, HR) associated with an increase in membrane conductance. There was no observed HR during a 10-min recording before insemination. The first HR had the following characteristics: peak potential, -72 ± 8 mV (mean $\pm$ s.d., $n = 53$); duration, 14 ± 3 s; membrane resistance near the peak, 28 ± 9 M Ω . The membrane conductance became 4.5 ± 1.8 times larger than that just before insemination. Successive HRs differed in minor ways: they became slightly longer and the peak became slightly less negative with a smaller increase in conductance (Fig. 1A). Summation of HRs was never observed, but they appeared in an all-or-none manner. The repetitive HRs persisted at almost constant intervals for at least 15 min while they were recorded, during which time some additional sperm attached to the egg after the initial rush of sperm. The resting membrane potential shifted gradually in the hyperpolarizing direction to ~ -40 mV, and the conductance increased while the I – V curve remained linear. A small depolarization which appeared immediately after insemination was not induced by sperm, but was caused by serum and/or adrenal gland extract in the sperm incubation medium (Fig. 1Ba, b).

In the conditions used in our experiments, fertilization did not occur in 30% of impaled eggs. Table 1 indicates that the occurrence of HR(s) was associated with sperm entry into an egg

Table 1 Correlation between HR(s) and fertilization

	HR (+)*	HR (–)
Sperm entry (+)*	44 eggs	4†
Sperm entry (–)	3‡	20

Sperm entry was determined by observation of male pronucleus or decondensation of sperm head.

* Occurrence of HR or sperm entry, irrespective of the number.

† Additional sperm attached to the egg after the electrode was removed. It is likely that these sperm entered the egg.

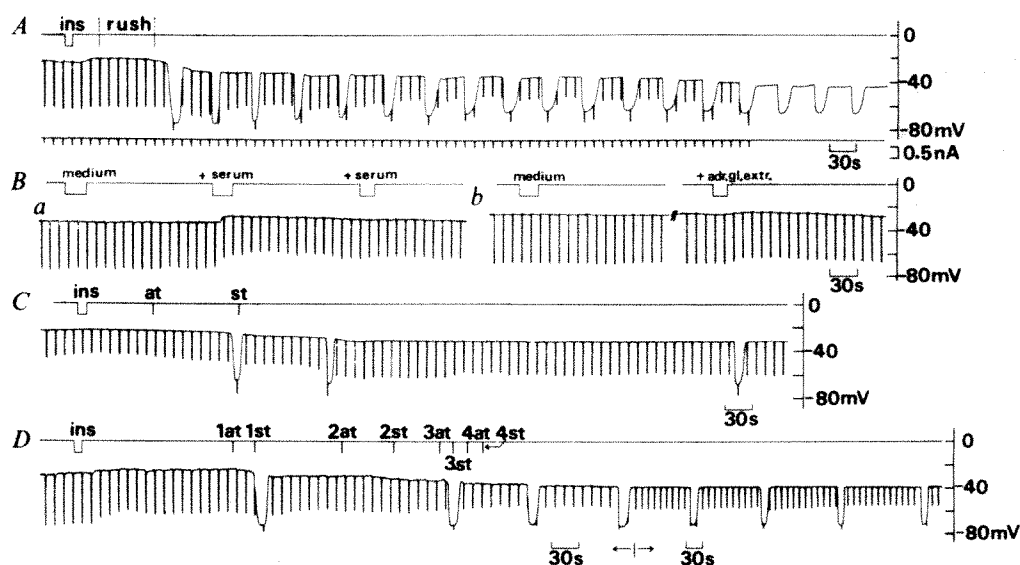
‡ The eggs had been damaged, thus decondensation of the sperm head might be prevented.

and that, conversely, sperm failed to enter eggs that had not shown any HR. When sperm with insufficient pretreatment were used, no HR was induced; these sperm have intact acrosomes and stick to the surface but never fuse with the egg¹⁰.

To correlate the occurrence of HR(s) with the interaction between sperm and egg, the number of sperm reaching an egg was restricted by reducing the concentration (1,000–2,000 sperm per 0.4 ml) and by placing a glass half-ring around the egg as a barrier. The first sperm reached and attached to the egg several minutes after insemination ('at' in Fig. 1). Then, using a micromanipulator, we added another half-ring to form a closed ring. Active flagellar motion decreased fairly suddenly, 10–90 s after the sperm had attached. The sperm then stopped moving completely in 10–15 s, becoming straight ('st' in Fig. 1). Table 2 shows that the first HR began in most cases within 10 s before or after the stop of the first sperm: there was a fairly good correlation between the two events.

In Fig. 1C, only one sperm bound to the egg throughout the experiment (Fig. 2A), and it was incorporated into the egg to form a male pronucleus (Fig. 2B). Three HRs appeared during 11 min after the sperm had stopped moving. In 10 out of 13 eggs penetrated by a single sperm, 2–4 HRs appeared within 2 min after the sperm had stopped. The next HR tended to occur after a pause of several minutes, as in Fig. 1C. Another three eggs behaved differently: one showed a continuous series of small HRs and two showed only one. The recurring HRs were also recorded in a zona-intact egg. In the case of Fig. 1D, four sperm attached and stopped separately from each other, and all were found to be incorporated into the egg. The stops of flagellar motion of additional sperm did not always coincide with the occurrence of a HR. However, multiple sperm entries caused a

Fig. 1 Fertilization potential induced by 16(A), one (C) and four sperm (D). To monitor a change in the membrane conductance, repetitive current pulses were applied through the intracellular electrode by means of a bridge circuit, with a current strength of 0.3 nA in A (bottom trace), and 0.2 , 0.5 , 0.5 and 0.15 nA in Ba to D, respectively. The upper trace shows events which were marked electrically by the observer, and also indicates zero potential: ins, insemination; rush, rush of sperm to attach to the egg surface; 1–4 at and st, attachment and stop of the first to fourth sperm. In B, a small depolarization associated with conductance increase was induced by addition of 20 μ l of sperm incubation medium containing 20% human serum (Ba) or 1% bovine adrenal gland extract (Bb). The incubation medium itself had no effect. The egg was insensitive to the second application of serum (Ba). There was a 3-min interruption in the middle of Bb. Concentrations of major ions in the egg medium were as follows (mM): Na $^+$, 119; K $^+$, 5.9; Ca $^{2+}$, 1.7; Mg $^{2+}$, 1.2; Cl $^-$, 102; HCO $_3^-$, 26 (equilibrated with CO $_2$ at pH 7.4).



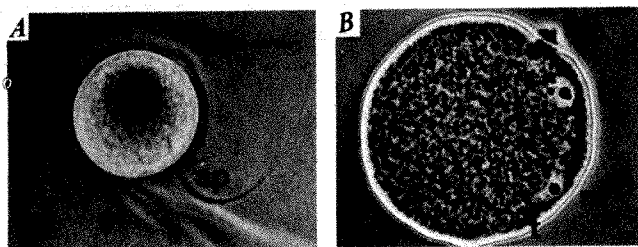


Fig. 2 The egg from which the record of Fig. 1C was obtained. **A**, After recording of potentials; **B**, after incubation for 4 h. The egg was pressed with a cover slip in **B** to visualize pronuclei. Scale bar for **A** and **B**, 50 μ m. sp, Sperm; el, electrode; pb, second polar body; f, female pronucleus; m, male pronucleus; t, sperm tail.

continuous series of HRs without any long pause. The series of HRs persisted although no more sperm became attached.

We conclude that the fertilization potential in hamster eggs consists of recurring hyperpolarizations. In our experiments, the time of recording the potentials after insemination corresponded to the period from contact between sperm and egg to, at most, the incorporation of the entire sperm head into the egg cytoplasm^{10,11}. The HR first appeared at about the time of the first sperm stop, when sperm and egg are likely to fuse¹⁰. The periodic appearance of HRs thereafter may be a consequence of a pre-existing ability of the egg, because similarly shaped HRs appear at almost constant intervals even without correlation with each sperm-egg fusion, as shown in the case of polyspermy (Fig. 1A, D). Repetitive hyperpolarizing responses each with durations of several to 20 s have been found in fibroblastic L cells^{12,13}, macrophages^{14,15} and sympathetic ganglion cells¹⁶. These HRs are due to an increase in the K permeability^{12,15} caused by an increase in intracellular Ca^{2+} ($[Ca^{2+}]_i$)^{14,16,17}. The HRs in hamster eggs had a reversal potential between -85 and -80 mV (our unpublished results), suggesting an increase in K permeability. The HRs in the above cells occur either spontaneously or by electrical, chemical or mechanical stimulations, whereas those in hamster eggs were induced neither by injecting current up to 5 nA nor by pushing the egg with a glass capillary, but they were induced by sperm about to be incorporated into the egg. It has been shown in sea urchin eggs^{18,19} that one of the early events on fertilization is the release of Ca^{2+} from intracellular stores and the consequent increase in $[Ca^{2+}]_i$ mediates several changes such as the discharge of the cortical granules. The HR in hamster eggs may reflect a cyclic change in $[Ca^{2+}]_i$ during incorporation of the sperm into the egg, although the functional significance is unknown.

In echinoderm and echiuran eggs, the membrane potential shifts to positive values during the early phase of the fertilization potential, due to an increase in membrane permeability to Na^+ and possibly Ca^{2+} (ref. 1). In the medaka fish, *Oryzias latipes*, fertilization triggers a depolarization of a few mV followed by a longer hyperpolarizing phase with an amplitude of ~30 mV, when the resting potential is more negative than -20 mV (ref. 20). In hamster eggs we could not record any depolarization preceding the HR in recording conditions that were able to detect a change in potential of at least 1 mV. This was also the case when the membrane potential was preliminarily shifted to

-64 to -68 mV by applying current. In one of four such cases, the membrane potential and effective resistance were -65 mV and 310 M Ω , respectively. We have proposed an equivalent electrical circuit for the egg plasma membrane²¹ analogous to circuits for subsynaptic membranes of nerve cells²². If one assumes a possible 1-mV depolarization caused by a shunting conductance (g_s) with a reversal potential of -20 mV, the g_s in the above egg is calculated as 72 pS. If the potential is clamped at -65 mV, the g_s causes a membrane current of 3.2 pA. Thus, the depolarizing response, if any, is negligibly small. Furthermore, preliminary experiments indicated that the potential-dependent fertilization block shown in echinoderm^{2,3}, echiuran⁴ and frog⁵ eggs does not exist in the hamster egg. These differences stimulate further experimentation in mammalian eggs.

We thank Drs A. Hanada and Y. Hirao for teaching us the techniques of fertilization *in vitro* and for valuable advice, Dr S. Ozawa for criticism, and Miss Y. Hodota for technical assistance.

Received 8 December 1980; accepted 3 February 1981.

- Hagiwara, S. & Jaffe, L. A. *Rev. Biophys. Bioengng* **8**, 385-416 (1979).
- Jaffe, L. A. *Nature* **261**, 68-71 (1976).
- Miyazaki, S. & Hirai, S. *Dev. Biol.* **70**, 327-340 (1979).
- Gould-Somero, M., Jaffe, L. A. & Holland, L. J. *Cell Biol.* **82**, 426-440 (1979).
- Cross, N. L. & Elinson, R. P. *Dev. Biol.* **75**, 187-198 (1980).
- Yanagimachi, R. *J. Reprod. Fert.* **18**, 275-286 (1969).
- Toyoda, Y. & Chang, M. C. *J. Reprod. Fert.* **36**, 9-22 (1974).
- Yanagimachi, R. *Biol. Reprod.* **3**, 147-153 (1970).
- Bavister, B. D., Yanagimachi, R. & Teichman, R. J. *Biol. Reprod.* **14**, 219-221 (1976).
- Yanagimachi, R. in *Current Topics in Developmental Biology* Vol. 12 (ed. Monroy, A.) 83-105 (Academic, New York, 1978).
- Yanagimachi, R. & Noda, Y. D. *Am. J. Anat.* **128**, 429-462 (1970).
- Nelson, P. G., Peacock, J. & Minna, J. J. *gen. Physiol.* **60**, 58-71 (1972).
- Okada, Y., Doida, Y., Roy, G., Tsuchiya, W., Inoue, K. & Inoue, A. *J. Membrane Biol.* **35**, 319-335 (1977).
- Gallin, E. K., Wiederhold, M. L., Lipsky, P. E. & Rosenthal, A. S. *J. cell. Physiol.* **86**, 653-662 (1975).
- Dos Reis, G. A. & Oliveira-Castro, G. M. *Biochim. biophys. Acta* **469**, 257-263 (1977).
- Kuba, K. & Nishi, S. *J. Neurophysiol.* **39**, 547-563 (1976).
- Okada, Y., Tsuchiya, W. & Inoue, A. *J. Membrane Biol.* **47**, 357-376 (1979).
- Steinhardt, R. A. & Epel, D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1915-1919 (1974).
- Steinhardt, R., Zucker, R. & Schatten, G. *Dev. Biol.* **58**, 185-196 (1977).
- Nuccitelli, R. *Dev. Biol.* **76**, 483-498 (1980).
- Miyazaki, S. *Dev. Biol.* **70**, 341-354 (1979).
- Eccles, J. C. *The Physiology of Synapses* (Springer, New York, 1964).

Electrical coupling between fibre cells in amphibian and cephalopod lenses

T. J. C. Jacob & G. Duncan

School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

The lenses of vertebrate and cephalopod eyes differ ontogenetically and in other respects. The vertebrate lens, derived from a single cell type, consists mainly of long fibre cells continuously produced by division and elongation of columnar epithelial cells near the lens equator. Almost 50% of the fibre cell surface consists of junctional complexes^{1,2} and the internal resistance, from point to point within the lens, is low compared with the surface membrane resistance³. Thus the vertebrate lens is expected to behave as a well coupled syncytial system⁴. The cephalopod lens, however, is formed by the fusion of two distinct cell types⁵; the anterior segment has the same ontogenetic origin as the cornea but the posterior segment shares a common origin with the retina, and the plane of contact of the two cell types can be seen in light-microscope sections⁵. Most of the lens is composed of long fibre cells similar in appearance to those found in the vertebrate lens, and membrane junctional regions between adjacent fibres have also been tentatively identified⁶. We now describe electrophysiological investigations of cellular communication in the cephalopod lens, which show marked differences in the intercellular electrical coupling within the vertebrate (amphibian) and cephalopod lens.

Table 2 Timing of the initiation of the first HR in relation to the first sperm 'stop'

Time (s)	Eggs	Time (s)	Eggs
-5-0	11	0-5	10
-10--5	5	5-10	6
-15--10	1	10-15	2
-----15	0	15-20	2
		20----	4

The time when the flagellar motion of the first sperm ceased was taken as zero time.

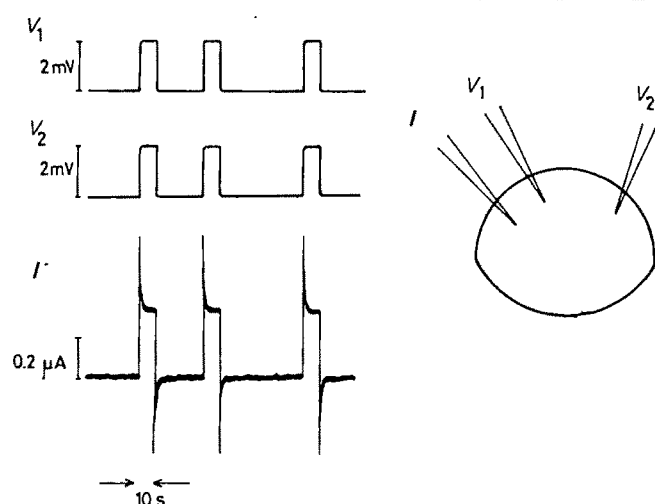


Fig. 1 Voltage-clamp responses at two positions (V_1 and V_2) within the cortex of the perfused lens of the frog *Rana pipiens* (for dissection and solution details, see ref. 9). The experimental arrangement allowed for two separate voltage-measuring circuits. The first electrode (V_1) measured the lens transmembrane voltage with respect to a reference in the bath and was coupled by a voltage-clamp circuit to the current-injection microelectrode (I) so that the lens could be clamped at any given voltage (see ref. 9 for further details). The voltage was also measured by an independent differential amplifier circuit between a third internal microelectrode (V_2) and a separate reference electrode in the solution flowing round the lens. The upper trace shows 2-mV command pulses superimposed on a resting voltage of -75 mV. The middle trace gives the response of the independent electrode where the resting baseline was -74 mV. V_2 in this case was placed one half of a lens diameter from V_1 . The clamp current (I) consists of a sharp spike, indicating current flow through the large membrane capacitance¹⁰ followed by a decline to a plateau value when current is flowing through the pure resistive components of the lens membranes. The voltage and current responses were found to be independent of the relative positions of the three widely-spaced electrodes.

Insertion of three microelectrodes into the frog lens does not significantly impair its membrane characteristics as stable, high (mean $\pm$ s.e.m. = 77.8 ± 0.7 mV from 29 lenses) resting potentials were obtained during these experiments. The voltage recorded at electrode V_1 was always within 1–2 mV of that recorded at V_2 . When V_1 was clamped at the resting voltage (Fig. 1), V_2 also remained clamped and when a 2-mV pulse (either hyperpolarizing or depolarizing) was superimposed on the resting voltage (V_1), a very similar change in baseline was also recorded at V_2 . The coupling ratio (V_1/V_2) obtained in four such experiments was 0.98 ± 0.1 .

However, the voltage-clamp responses obtained from three similar internal electrode experiments on the perfused lens of the cephalopod *Sepiolo atlantica*, depended on the relative positions of the current and voltage electrodes (Fig. 2). When the current and command voltage electrodes were both placed in the anterior segment and the independent voltage electrode in the posterior, a 2-mV command pulse was not accompanied by a 2-mV response at the posterior electrode (Fig. 2a); the coupling ratio defined by V_p/V_a was in fact only 0.06. These responses were mirrored when the command and current electrodes were

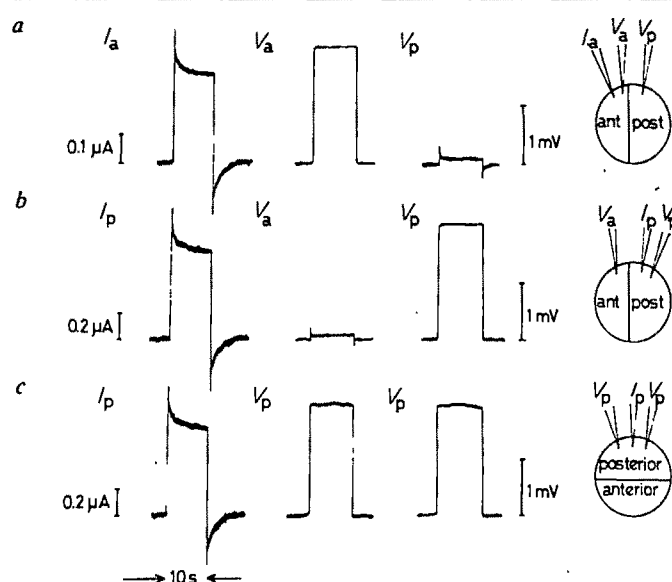


Fig. 2 Voltage-clamp responses at the anterior and posterior segments of the perfused cephalopod lens (for details of dissection methods and solution composition, see ref. 6). The command voltage microelectrode and current injection electrode were always placed in the same segment of the lens and the shape of the voltage and current signals obtained in response to a command pulse of 2 mV were very similar to those obtained from the vertebrate lens (Fig. 1). However, the magnitude and time course of the response measured by the independent voltage circuit depended on whether the command and independent electrode were located in the same lens segment. When they were in different segments (as in a and b), the independent signal was greatly reduced in amplitude and its shape followed that of the clamp current rather than the command voltage. When the two voltage electrodes were in the same segment (as in c), the amplitudes and shapes of the responses were very similar. Note the difference in vertical scale magnitudes of the clamp currents in a and b, which is due to the fact that the resistance of the anterior segment is greater than that of the posterior (Table 1).

both in the posterior segment, with the independent electrode in the anterior (Fig. 2b); the coupling ratio (V_a/V_p in this case) was again 0.06 and the mean value obtained from eight such experiments was 0.072 ± 0.015 . When the three electrodes were placed in the same segment (either anterior or posterior), the command pulse was followed precisely by the independent electrode (Fig. 2c). Hence cells within one segment are closely coupled, but there is little communication between segments.

The resting voltages recorded in the two segments were very similar (Table 1), but the resistances were quite different. However, the surface area of the anterior segment in the *Sepiolo* lens is about half that of the posterior (our unpublished observations), thus when the data are expressed as specific resistance ($k\Omega \text{ cm}^2$) the two values are almost equal. This suggests that although the membranes of the fibre cells in the two segments have a different origin and although they are not coupled, their membrane permeability properties may be quite similar. The specific resistances of the *Sepiolo* segments are both considerably smaller than the specific resistance of the frog lens, which Delamere and Duncan⁶, ascribed to the very high potassium permeability of the *Sepiolo* lens membranes.

Table 1 Accumulated potential and resistance data for the amphibian (*Rana pipiens*) and cephalopod (*Sepiolo atlantica*) lenses

	Potential (mV)	Resistance (k Ω)	Surface area ($\times 10^{-2} \text{ cm}^2$)	Specific resistance (k $\Omega \text{ cm}^2$)
<i>Rana</i> lens	-77.8 ± 0.7 (29)	10.9 ± 0.48	33.9 ± 1.17	3.69 ± 0.144
<i>Sepiolo</i> lens anterior	-62.5 ± 1.4 (15)	6.34 ± 0.49	6.9 ± 0.3	0.44 ± 0.026
posterior	-62.2 ± 2.0 (9)	3.49 ± 0.87	15.3 ± 0.59	0.53 ± 0.07

Data are mean $\pm$ s.e.m. and values in parentheses represent the number of lenses in each case.

The fact that the two segments of the cephalopod lens are not closely coupled (Fig. 2a, b) poses interesting problems for the regulation of growth and raises a further issue concerning the crystallin protein composition of the two halves. As they are of different ontogenetic origin, and because they remain functionally separate throughout development, it is very likely that they are composed of different protein species. Although the proteins of cephalopod lenses have been investigated in some detail⁷ and the nuclear and cortical components seem to be different,⁸ no studies have been made of the two halves separately.

The cephalopod lens thus provides a unique opportunity for comparative evolutionary and developmental studies. Not only has it developed one mechanism in parallel with the vertebrate lens for producing a highly refractile, transparent focusing system (that is, by close-packed array of fibre cells of surface ectodermal origin), but it has also evolved a second and apparently independent mechanism for producing crystallins and fibre cells from a second cell type (vesicular epithelium).

Received 2 December 1980; accepted 29 January 1981.

1. Philipson, B. T., Hanninen, L. & Balazs, E. A. *Expl Eye Res.* **21**, 205–219 (1975).
2. Kuszak, J., Maisel, H. & Harding, C. V. *Expl Eye Res.* **27**, 495–498 (1978).
3. Duncan, G. *Expl Eye Res.* **8**, 406–414 (1969).
4. Duncan, G. in *The Eye* Vol. 5 (eds Davson, H. & Graham, L. T.) 357–398 (Academic, London 1974).
5. Duke-Elder, S. *System of Ophthalmology* Vol. 1 (Kimpton, London, 1958).
6. Delamere, N. A. & Duncan, G. *J. Physiol., Lond.* **272**, 167–186 (1977).
7. Clayton, R. M. in *The Eye* Vol. 5 (eds Davson, H. & Graham, L. T.) 399–494 (Academic, London, 1974).
8. Dohrn, A. *Expl Eye Res.* **9**, 297–299 (1970).
9. Jacob, T. J. C. & Duncan, G. *Expl Eye Res.* **31**, 505–512 (1980).
10. Duncan, G., Patmore, L. & Pynsent, P. B. *J. Physiol., Lond.* **312** (in the press).

Rhodopsin-like photosensitivity of isolated chicken pineal gland

Takeo Deguchi

Department of Medical Chemistry, Tokyo Metropolitan Institute for Neurosciences, 2–6 Musashidai, Fuchu-city, Tokyo 183, Japan

The pineal gland of lower vertebrates such as fish, amphibians and reptiles is primarily a photoreceptor organ and its electrical activity changes in response to environmental lighting^{1–3}, whereas the mammalian pineal gland is a secretory organ and no longer responds to direct illumination. It is uncertain whether the avian pineal gland retains a photoreceptive capability^{4–8}, and also whether the gland is a photosensor involved in the photo-periodic control of reproductive systems and circadian rhythms in birds^{9–14}. Recently, I and other workers have shown that the circadian rhythm of serotonin *N*-acetyltransferase activity persisted in organ cultures of chicken pineals, indicating the presence of an endogenous circadian oscillator in the gland^{15–19}. Also, the observation that direct illumination of cultured chicken pineals suppressed the night-time increase of *N*-acetyltransferase activity^{16,20} suggests that the gland contains a photoreceptor. I report here that the action spectrum of the photosensitivity of the isolated chicken pineal resembles the absorption spectrum of rhodopsin.

Chickens (White Leghorn cockerels) were maintained in our animal facility with light from 7 a.m. to 7 p.m. for 14–18 days after hatching. The chickens were killed by decapitation at 6 p.m. and five pineals were placed on nylon mesh in a Falcon plastic organ culture dish (no. 3037) containing 1.5 ml of BGJb medium without phenol red. They were cultured as described elsewhere¹⁶ from 7 p.m. to 1 a.m. under a monochromatic and energy-calibrated photic stimulus. As controls, a group of pineals was cultured in darkness (dark control) and another

Table 1 Effect of depolarizing and hyperpolarizing agents on *N*-acetyltransferase rhythm in organ cultures of chicken pineal gland

Addition	<i>N</i> -acetyltransferase (pmol per 10 min per pineal)	
	Light	Dark
None	152 ± 12	997 ± 50
55 mM KCl	434 ± 58	990 ± 131
Nonactin (20 µM)	115 ± 13	107 ± 15
Valinomycin (4 µM)	129 ± 10	135 ± 11
Before culture	98 ± 6	

Pineals were organ-cultured from 7 p.m. to 1 a.m. in darkness or under 70 µW cm⁻² of white light in the presence of the agents indicated. There were five samples in each group and the results are the means ± s.e.m.

group under 70 µW cm⁻² of white light from a xenon arc lamp (light control).

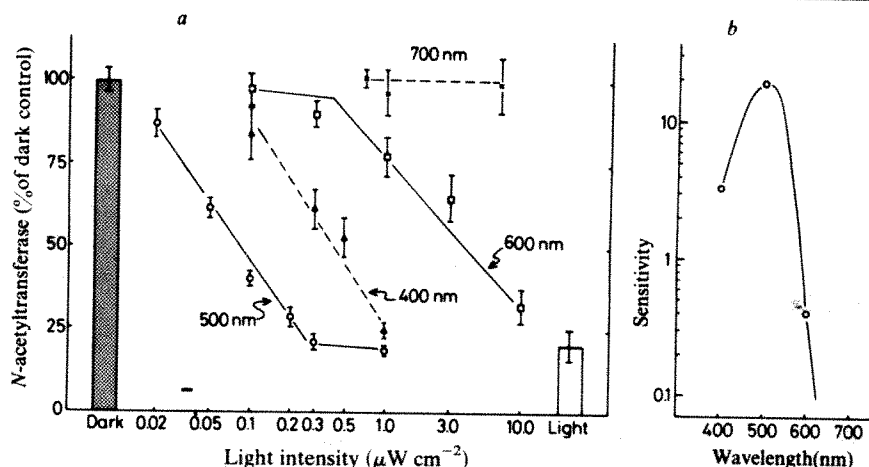
When pineals were cultured from 7 p.m. to 1 a.m. in darkness, *N*-acetyltransferase activity increased from 75–120 to 750–1,300 pmol per 10 min per pineal. When they were cultured under continuous illumination, the enzyme activity increased only to 150–250 pmol per 10 min per pineal. Culture under various intensities of 400, 500, 600 and 700 nm wavelengths (Fig. 1a) showed 500 nm to be the most effective wavelength for inhibiting *N*-acetyltransferase activity. At this wavelength there was a significant decrease of the enzyme activity with an intensity of 0.02 µW cm⁻² ($P < 0.05$), and maximum inhibition with 0.3 µW cm⁻², which was equivalent to that caused by 70 µW cm⁻² of white light. With the wavelength of 400 nm maximum inhibition occurred at 1 µW cm⁻², while 600 nm was far less effective, showing no significant inhibition with 0.3 µW cm⁻² and a submaximum inhibition with 10 µW cm⁻². The wavelength of 700 nm showed no effect on enzyme activity with 7 µW cm⁻² or less. When the reciprocals of the light intensities producing 50% inhibition of the night-time increase of *N*-acetyltransferase activity were plotted as sensitivities on the ordinate in a logarithmic scale (Fig. 1b), the sensitivity curve resembled the absorption spectrum of rhodopsin²¹. On examining the effects of wavelengths from 400 to 800 nm at 0.3 µW cm⁻², maximum inhibition occurred at 500 nm and a submaximum inhibition at other wavelengths (Fig. 2). The most effective wavelengths were 500 and 520 nm, and there was a sharp cutoff of the inhibitory effect at longer wavelengths, with a gradual decrease at shorter wavelengths.

A high enough concentration of KCl (55 mM) to depolarize the cell membrane partially restored the light-induced suppression, but did not enhance the dark-time increase of *N*-acetyltransferase activity (Table 1). On the other hand, the dark-time increase of *N*-acetyltransferase activity was completely inhibited by potassium ionophores (nonactin or valinomycin), which hyperpolarize the cell membrane. I have previously shown that cocaine, which prevents depolarization of neural cells, and catecholamines, which induce hyperpolarization in the rat pinealocyte, block the night-time elevation of *N*-acetyltransferase activity²².

The threshold light intensity of the chicken pineal is in the range of the threshold intensities of several photosensitive systems²³. As dispersed cells responded to a photic stimulation in the same manner as organ culture²⁴, it seems reasonable to speculate that a rhodopsin-like photosensitive substance is incorporated in the cell membrane of pinealocytes. Electron microscopic studies of avian pineals revealed cells with whorls of lamellae and a modified structure, which some investigators consider photosensory^{6,7} but others believe to be abortive or vestigial photoreceptor cells⁸. Such a structure might participate in the regulation of *N*-acetyltransferase activity.

It is firmly established that the membrane potential of retinal photoreceptor cells of vertebrates and of frog pineals are depolarized in darkness and hyperpolarized by photic stimulation^{3,25}. It is therefore conceivable that direct illumination of the

Fig. 1 Effect of various intensities of monochromatic light on the night-time increase of *N*-acetyltransferase activity in organ cultures of chicken pineal. A monochromatic stimulator (model TD-3, Sanso Seisaku) with a xenon arc lamp (XW-500, 500 W) was used. White light from the xenon lamp was divided into three channels and the light energy was calibrated by using an optical wedge and neutral density filter. Monochromatic light was obtained by interference filters (half width <11 nm) and the beam was transmitted into a CO₂ incubator through a quartz window. The stimulating light was reflected downwards with a mirror and projected on to a culture dish. The intensity of light was measured at the level of the dish by use of a radiometer (model 470D, Sanso Seisaku) and expressed as $\mu\text{W cm}^{-2}$. *a*, Pineals were cultured from 7 p.m. to 1 a.m. under the intensities of monochromatic light indicated. *N*-acetyltransferase activity is shown as per cent of dark control, which was run in each experiment. Points and bars show means and s.e.m. of values of five samples. *b*, Reciprocals of the intensities that produced 50% inhibition are plotted as sensitivities on the ordinate in a logarithmic scale.



chicken pineal produces a slow, long-lasting hyperpolarization of the pinealocyte membrane and prevents the nocturnal increase of *N*-acetyltransferase activity. Such speculation is supported by the present finding that depolarizing agents restored the light-suppression of *N*-acetyltransferase activity and hyperpolarizing agents prevented the night-time increase of the enzyme activity.

Kato *et al.* reported that red, but not green, radioluminous paint applied to the skin above the pineal gland of Japanese quail promoted testicular development¹⁴. Woodard *et al.* also showed that in Japanese quail red light accelerated sexual maturity more effectively than green light²⁶. In contrast, Menaker showed that green light (0.1 lx) was effective in entraining the circadian rhythm of perching activity in blinded sparrows²⁷. In my study, green light was most effective in inhibiting the *N*-acetyltransferase rhythm. The discrepancy between these observations is unclear. However, in the present study pineals were directly illuminated, whereas in other studies animals were placed under a specified wavelength of light. It is possible that shorter wavelengths were absorbed by skin, skull, blood and other tissues before reaching the pineals. Alternatively, the hypothalamic, but not pineal, photoreceptor may sense longer wavelengths of light, thus affecting photoreproductive function in living birds.

Menaker and his associates have demonstrated that in house sparrows the pineal gland functions as a biological clock that

controls the circadian rhythms of locomotor activity and body temperature²⁸. The gland presumably regulates such functions by secretion of melatonin, a pineal hormone^{29,30}. The activity of serotonin *N*-acetyltransferase, the key enzyme in the synthesis of melatonin³¹⁻³³, shows a circadian rhythm and responds to direct illumination in the isolated chicken pineal glands, as reported here. Recently, Takahashi *et al.*³⁴ have shown that the synthesis and release of melatonin from cultured chicken pineal glands also exhibit a circadian change and respond to environmental lighting, suggesting that the light perception of the avian pineal gland, by affecting the synthesis of melatonin, has a physiological role in the regulation of functions and behaviour in birds.

This research was in part supported by a Grand-in-Aid for Special Project Research on Animal Behaviors and by a Grant-in-Aid for Scientific Research (548121) from the Ministry of Education of Japan.

Received 24 October 1980; accepted 11 February 1981.

- Wurtman, R. J., Axelrod, J. & Kelly, D. E. in *The Pineal*, 9-39 (Academic, New York, 1968).
- Dodt, E. in *Handbook of Sensory Physiology* Vol. VII/3B (ed. Jung, R.) 113-140 (Springer, Berlin, 1973).
- Morita, Y. in *Brain-Endocrine Interaction* (eds Knigge, K. M., Scott, D. E., Kobayashi, H. & Ishii, S.) 346-387 (Karger, Basel, 1975).
- Morita, Y. *Experientia* **22**, 402-403 (1966).
- Ralph, C. L. & Dawson, D. C. *Experientia* **24**, 147 (1968).
- Quay, W. B., Renzoni, A. & Eakin, R. M. *Rev. Biol.* **61**, 371-393 (1968).
- Bischoff, M. B. *J. ultrastruct. Res.* **28**, 16-26 (1969).
- Oksche, A. & Vaupel-Von Harnack, M. Z. *Zellforsch.* **69**, 41-60 (1966).
- Harrison, P. C. & Becker, W. C. *Proc. Soc. exp. Biol. Med.* **132**, 161-164 (1969).
- Homma, K. & Sakakibara, Y. in *Biochronometry* (ed. Menaker, M.) 333-341 (National Academy of Science, Washington DC, 1971).
- Menaker, M., Roberts, R., Elliott, J. & Underwood, H. *Proc. natn. Acad. Sci. U.S.A.* **67**, 320-325 (1970).
- Oksche, A., Kirschstein, H., Kobayashi, H. & Farner, D. A. Z. *Zellforsch. mikrosk. Anat.* **24**, 247-274 (1972).
- Hisano, N., Cardinali, D. P., Rosner, J. M., Nagle, C. A. & Tramezzani, J. H. *Endocrinology* **91**, 1318-1322 (1972).
- Kato, M., Kato, T. & Oishi, T. *Proc. Jap. Acad.* **43**, 220-223 (1967).
- Deguchi, T. in *Integrative Control Functions of the Brain* Vol. 1 (eds Ito, M., Tsukahara, N., Kubota, K. & Yagi, K.) 345-347 (Kodansha, Tokyo/Elsevier, Amsterdam, 1978).
- Deguchi, T. *Science* **203**, 1245-1247 (1979).
- Binkley, S. A., Riebelman, J. B. & Reilly, K. B. *Science* **202**, 1198-1201 (1978).
- Kasal, C. A., Menaker, M. & Perez-Polo, J. R. *Science* **203**, 656-658 (1979).
- Wainwright, S. D. & Wainwright, L. K. *Can. J. Biochem.* **57**, 700-709 (1979).
- Wainwright, S. D. & Wainwright, L. K. *J. Neurochem.* **35**, 451-457 (1980).
- Wald, G. & Brown, P. K. *Nature* **177**, 174-176 (1956).
- Deguchi, T. *J. Neurochem.* **33**, 45-51 (1979).
- Andresen, M. C. & Brown, A. M. *J. Physiol. Lond.* **287**, 267-282 (1979).
- Deguchi, T. *Nature* **282**, 94-96 (1979).
- Tomita, T. *Q. Rev. Biophys.* **3**, 179-222 (1970).
- Woodard, A. E., Moore, J. A. & Wilson, W. O. *Poultry Sci.* **47**, 1733-1737 (1967).
- Menaker, M. *Proc. natn. Acad. Sci. U.S.A.* **59**, 414-421 (1968).
- Menaker, M., Takahashi, J. S. & Eskin, A. A. *Rev. Physiol.* **40**, 501-526 (1978).
- Gwinner, E. & Benzinger, I. *J. comp. Physiol.* **127**, 209-213 (1978).
- Hendel, R. C. & Turek, F. W. *Physiol. Behav.* **21**, 275-278 (1978).
- Klein, D. C. & Weller, J. L. *Science* **169**, 1093-1095 (1970).
- Deguchi, T. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3547-3550 (1972).
- Axelrod, J. *Science* **184**, 1341-1348 (1974).
- Takahashi, J. S., Hamm, H. & Menaker, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2319-2322 (1980).

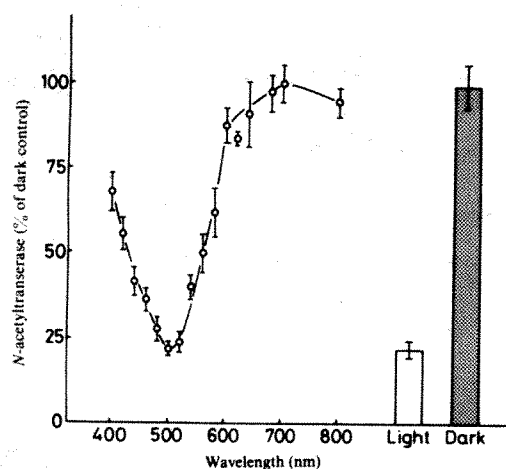


Fig. 2 Effect of various wavelengths on the *N*-acetyltransferase activity in organ culture. Pineals were cultured from 7 p.m. to 1 a.m. under an intensity of $0.3 \mu\text{W cm}^{-2}$ of the monochromatic light indicated. The results are shown as per cent of dark control. Points and bars indicate means and s.e.m. of five samples.

Functionally significant central-pair rotation in a primitive eukaryotic flagellum

Charlotte K. Omoto*† & George B. Witman*

* Department of Biology, Princeton University, Princeton, New Jersey 08544, USA

† Biology Division, California Institute of Technology, Pasadena, California 91125, USA

There is now considerable evidence that the basis for ciliary and flagellar movement is an active sliding between peripheral doublet microtubules which, when resisted by structures within the axoneme, leads to axonemal bend formation¹⁻⁴. In contrast, relatively little is known about the control mechanisms which coordinate the interdoublet sliding and axonemal bending to produce the effective motion observed in various cilia and flagella⁵. One component of the axoneme which may be involved in this control is the central pair of microtubules^{6,7}. To learn more about the action of the central pair, we have studied the tiny uniflagellate marine alga, *Micromonas pusilla*. The central tubules of the *M. pusilla* flagellum extend for several micrometres beyond the termination of the peripheral doublets⁸, thus permitting direct observation of the central pair during flagellar movement. Our findings, reported here, indicate that in living *M. pusilla* the central pair of microtubules undergoes continuous rotation in one direction. This rotation provides the motive force for the cell.

Our observations on the ultrastructure of the flagellum of *M. pusilla* (previously named *Chromulina pusilla*⁹) agree well with those of Manton⁸. The flagellum is ~5 µm long and has the appearance of a short thick stub terminated by a long thin fibre (Fig. 1a). The thicker proximal portion is about 0.2 µm in diameter and contains the normal '9+2' array of microtubules, but is less than 1 µm long (Fig. 1e). The thinner distal portion is ~70 nm in diameter and contains only the central pair of microtubules (Fig. 1b), which are ~5 µm in length and extend for ~4 µm beyond the termination of the peripheral doublets. For most of this distal portion, the flagellar membrane closely surrounds the central pair. Figure 1b and c show the region of transition between the '9+2' flagellum and the thin '0+2' portion containing only the central tubules. At their proximal end, the central tubules terminate just above the basal body (Fig. 1f), as in other cilia and flagella^{10,11}.

The flagella of living *M. pusilla* can be observed by dark-field light microscopy. In free-swimming cells, the distal '0+2' portion takes on a helical configuration and seems to rotate, pushing the cells through the medium like a propeller. This portion of the flagellum also appears to rotate in cells which are stuck to the glass slide by their cell bodies (Fig. 2). The short '9+2' region is usually obscured by glare from the cell body, but in fortuitous views it is seen to wave back and forth, as would be expected for a very short flagellum undergoing a normal beat cycle. The '0+2' portion of a non-moving flagellum retains its helical configuration. By focusing through such a flagellum, we determined that the helix is left-handed.

One cannot be sure from the above observations alone that the movement of the '0+2' portion of the flagellum is due to actual rotation as opposed to propagation of helical waves along a non-rotating flagellum. However, these two possibilities can be distinguished by anchoring the cell to the glass slide by its flagellum and observing the resulting motion of the cell body. When *M. pusilla* are placed on silicone-coated slides, many cells stick to the slide by one end of the cell body and spin rapidly around the point of attachment, as previously described by Manton and Parke⁹. However, some cells stick to the slide by the flagellum; when this occurs, the cell body rotates (Figs 3, 4). In

Fig. 3, the rotational axis is almost perpendicular to the slide, so that the rotation looks like that of a wheel viewed from the side. In Fig. 4, the rotational axis is nearly parallel to the glass slide and rotation is inferred by the changes seen in the shape and the markings of the cell. The position and curvature of the flagellum of this cell remain constant throughout the 360° rotation of the cell body, indicating that the flagellum is stuck to the slide and immobilized over most of its length. In all such cells so far observed, the cell bodies were rotating clockwise as viewed from the flagellum. This would correspond to an anticlockwise rotation of the '0+2' portion, looking from tip to base. These observations indicate that the distal portion of the *M. pusilla* flagellum, and by inference, the pair of central tubules therein,

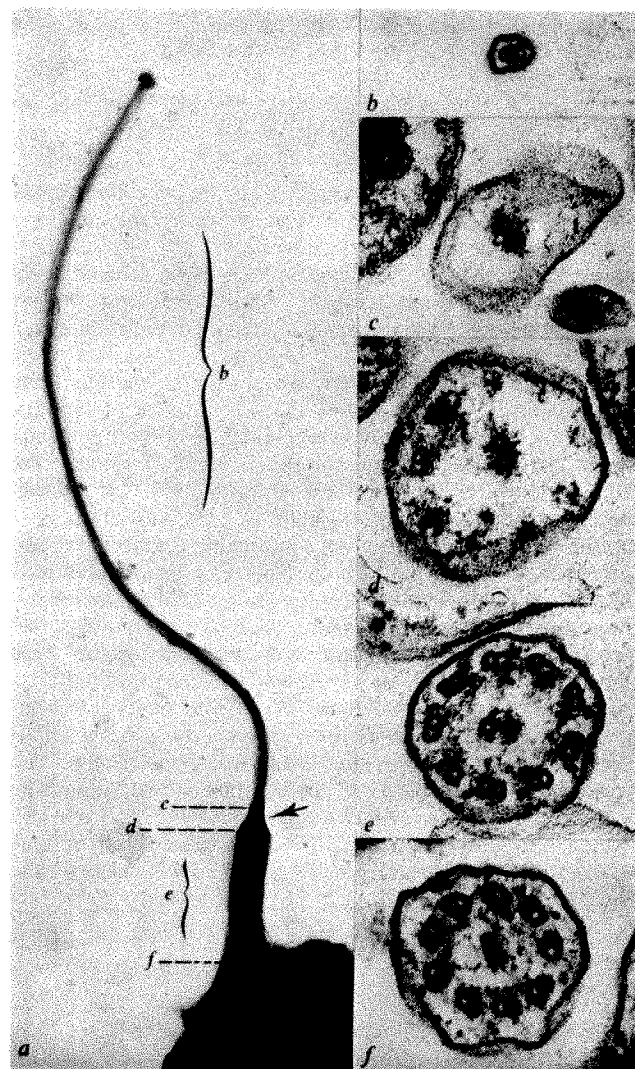


Fig. 1 Structure of *M. pusilla* flagellum. *a*, Electron micrograph of a negatively stained *M. pusilla*. The distal portion of the flagellum beyond the arrow is about 70 nm in diameter and is too thin to contain the usual '9+2' axoneme. The dotted lines and brackets labelled *b-f* indicate the approximate levels of the corresponding cross-sections. *M. pusilla* was grown under continuous light on Ergschieber's medium²⁰. A drop of the cell culture was placed on a Formvar-coated grid and stained with 2% uranyl acetate. $\times 15,275$. *b*, Cross-section through distal portion of the flagellum where only the central tubules remain. Note that the membrane closely surrounds the two tubules, which seem to be closer to each other than in the usual '9+2' structure. *c*, Cross-section through the flagellum just distal to the point where the peripheral tubules have terminated. *d*, Cross-section where some of the peripheral tubules have terminated. *e*, Cross-section through the proximal portion of the flagellum, showing the typical '9+2' axoneme. *f*, Cross-section at the base of the flagellum, showing the axosome just proximal to the origin of the central tubules. *b-f*, $\times 83,800$. *M. pusilla* was 'instantaneously fixed'¹² in glutaraldehyde- OsO_4 , embedded in agar, then dehydrated and embedded in Spurr's plastic²¹. Sections were stained with uranyl acetate and Reynold's lead citrate²².



Fig. 2 Sequence from 16-mm cine film of *M. pusilla*. The cell is stuck to the slide and the flagellum seems to rotate. Most of the flagellum seen is the '0+2' region, the '9+2' region being too short to be visible much beyond the glare of the cell body. Dark-field optics⁶; consecutive frames progress from left to right. 64 frames per s. $\times 2,000$.

rotates. The evidence does not rule out the possibility that the whole flagellum, including the '9+2' region, rotates, but this seems highly unlikely.

The anticlockwise rotation together with the left-handed twist of the flagellum are consistent with the observed progression of the cell through the medium, with the flagellum pushing the cell. Interestingly, the direction of central-pair rotation in *M. pusilla* is the same as that reported for the central pair in *Paramecium*^{7,12}.

The membrane which surrounds the central microtubules probably rotates with the central pair; this is especially evident in a rotating cell which is anchored to the slide by its flagellum, but even in free-swimming cells connections between the central pair and the membrane might cause the membrane to move with the underlying tubules. Such a movement would result in a shear zone in the flagellar membrane at some point along the flagellum, probably at the base of the '0+2' portion. We have observed no discontinuity in the cell or flagellar membrane of *M. pusilla*, so the surface shear must be readily absorbed by the fluid membrane. A similar situation occurs in the plasma membrane of some devescovinid flagellates, where one portion of the cell

surface continuously rotates relative to the rest of the cell surface¹³.

We do not know what causes the central pair to rotate. One possibility is that the rotation is a passive response to three-dimensional bending of the '9+2' portion of the flagellum. Alternatively, rotation may be the result of interactions between the radial spokes and the central sheath in the '9+2' region, or might be caused by a rotary motor in the axosome where the central tubules originate. Further studies will be necessary to determine the basis for the rotation.

Central-pair rotation has now been demonstrated for *M. pusilla* flagella, and implicated in *Paramecium* cilia^{7,12}, *Opalina* cilia¹⁴ and *Synura* flagella¹⁵. Central-pair rotation may therefore occur in the cilia and flagella of many organisms. In *M. pusilla*, rotation of the central pair propels the cell through its medium. The function of central-pair rotation in the normal '9+2' axoneme is unknown, but several observations suggest that the central microtubules are involved in the control of ciliary and flagellar movement. Flagella lacking the central tubules are paralysed, although the ability of the dynein arms to cause sliding is apparently normal¹⁶. The central tubule-central

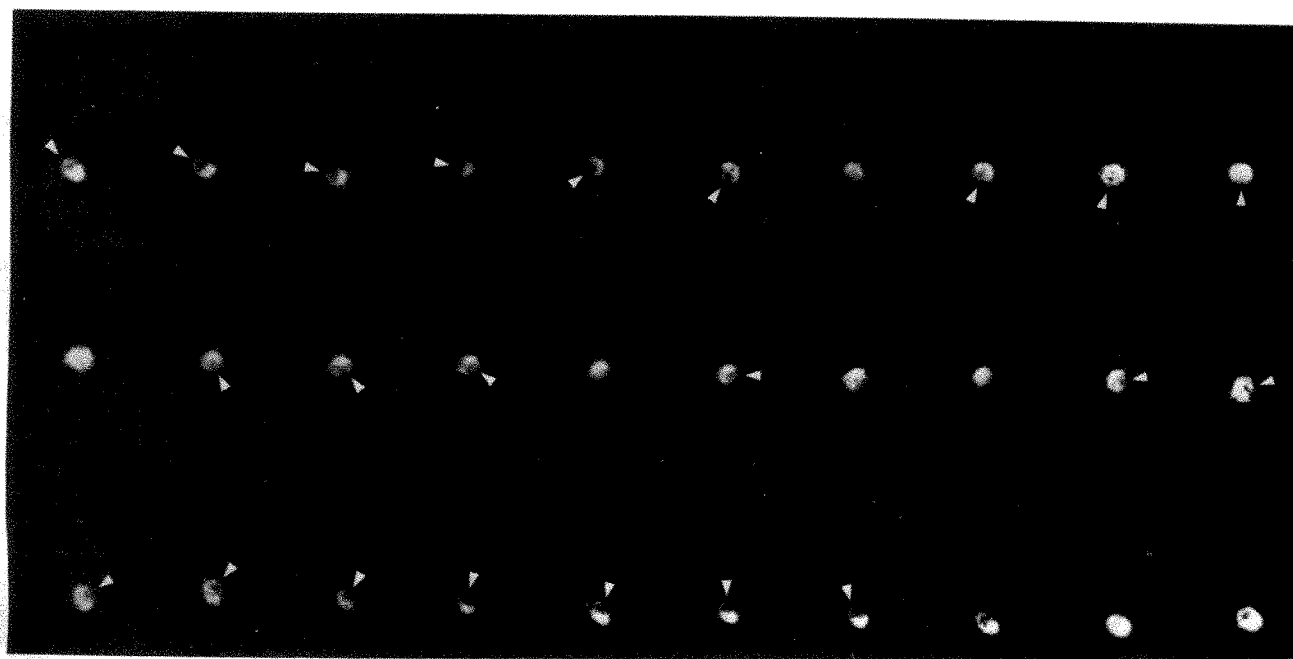


Fig. 3 Film sequence showing a cell stuck by its flagellum (visible in the original cine film) to a silicone-coated slide and the cell body rotating. The slide was coated with silicone by wiping it with silicone-treated lens paper (Sight Savers, Dow-Corning). The sequence was printed to enhance the visibility of the dark spot (white arrowheads) on the cell body. The cell-body rotation can be followed using the dark spot as a marker. The cell is rotating anticlockwise. Consecutive frames run from left to right beginning in the upper left-hand corner. Dark-field optics, 64 frames per s. $\times 2,880$.

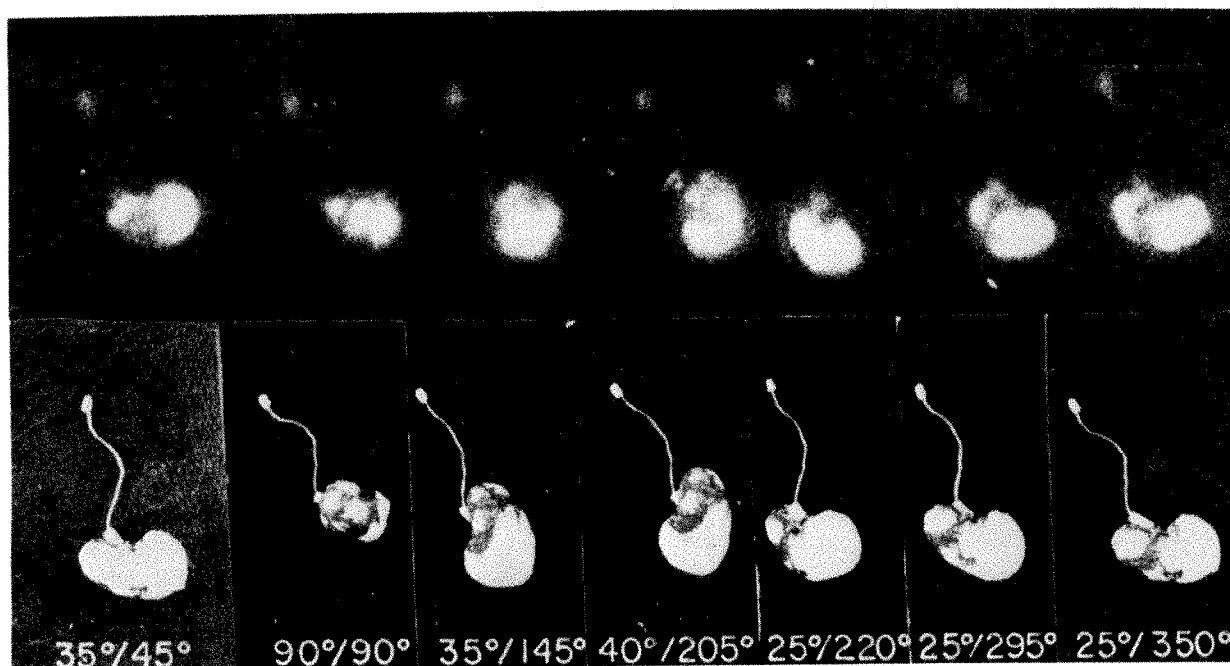


Fig. 4 Top row: dark-field light micrographs of *M. pusilla*. The cell is attached to a silicone-coated slide by its flagellum and the asymmetric cell body is rotating. Recorded on continuously moving film with stroboscopic illumination at 60 Hz. $\times 3,200$. Second row: photos of a styrofoam and wire model. The wire, representing the '0 + 2' portion, is held in place and the angle of insertion of the wire and the cell body orientation are changed and photographed. The numbers below each image indicate the angle of insertion relative to the cell surface, and the angle through which the cell body has rotated relative to the surface to which the flagellum is attached, respectively.

sheath complex is asymmetrical in several organisms, and this asymmetry may be related to the formation of asymmetrical axonemal bends^{6,16}. The orientation within the axoneme of the central pair of microtubules is correlated to beat direction in a variety of cilia and flagella^{14,17,18}; moreover, the orientation of the central pair has been reported to change along the length of the axoneme^{12,19}, with change in beat direction¹⁴, and during the beat cycle^{7,12,15}. Thus, it has been proposed that the central microtubules undergo a systematic movement correlated with the beat cycle and that this movement is involved in regulating the motion of the axoneme, possibly by determining which doublets undergo active sliding at any particular phase in the beat cycle⁷. Regulation of dynein arm activity around the axoneme would be particularly important in cilia and flagella having a three-dimensional waveform.

The cultures of *M. pusilla* were provided by the Culture Centre of Algae and Protozoa (Cambridge, UK) and the Culture Collection of Algae (University of Texas, Austin). We thank Jana Olson Kiefer for providing the growth medium, Nick Minervini for assistance with photography and electron microscopy, Gary Gorbysky for his suggestion of using 'Sight Savers' for coating slides and Drs Ching Kung, Howard Berg and Charles J. Brokaw for helpful criticism and encouragement. Figure 4 was made possible by the invaluable assistance of Dr Brokaw at the California Institute of Technology, Kerckhoff Marine Laboratory. This work was supported in part by NIH grants GM 21586 to G.B.W., GM 07445 to C.K.O. and GM 18711 to C.J.B.

Received 8 October; accepted 2 February 1981.

1. Satir, P. *J. Cell Biol.* **39**, 77-94 (1968).
2. Summers, K. E. & Gibbons, I. R. *Proc. natn. Acad. Sci. U.S.A.* **68**, 3092-3096 (1971).
3. Gibbons, B. H. & Gibbons, I. R. *Biochem. biophys. Res. Commun.* **73**, 1-6 (1976).
4. Shingyogi, C., Murakami, A. & Takahashi, K. *Nature* **265**, 269-270 (1977).
5. Brokaw, C. J. & Gibbons, I. R. in *Swimming and Flying in Nature* (eds Wu, T. Y.-T. et al.) 89-126 (Plenum, New York, 1975).
6. Bessen, M. et al. *J. Cell Biol.* **86**, 446-455 (1980).
7. Omoto, C. K. & Kung, C. *Nature* **279**, 532-534 (1979).
8. Manton, I. *J. mar. Biol. Ass. U.K.* **38**, 319-333 (1959).
9. Manton, I. & Parke, M. *J. mar. Biol. Ass. U.K.* **39**, 275-298 (1960).
10. Ringo, D. L. *J. Cell Biol.* **35**, 543-571 (1967).
11. Dute, R. & Kung, C. *J. Cell Biol.* **78**, 451-464 (1978).
12. Omoto, C. K. & Kung, C. *J. Cell Biol.* **87**, 33-46 (1980).
13. Tamm, S. L. & Tamm, S. J. *Cell Sci.* **20**, 619-639 (1976).
14. Tamm, S. L. & Horridge, G. A. *Proc. R. Soc. B* **175**, 219-233 (1970).
15. Jarosch, R. & Fuchs, B. *Protoplasma* **85**, 285-290 (1975).
16. Witman, G. B. et al. *J. Cell Biol.* **76**, 729-747 (1978).

17. Afzelius, B. A. *J. Biophys. Biochem. Cytol.* **9**, 383-394 (1961).
18. Gibbons, I. R. *J. Biophys. Biochem. Cytol.* **11**, 179-205 (1961).
19. Gibbons, I. R. in *Molecules & Cell Movement* (eds Inoué, S. & Stephens, R. E.) 207-232 (Raven, New York, 1975).
20. Starr, R. C. *J. Phycol.* **14**, Suppl., 47-100 (1978).
21. Spurr, A. R. *J. ultrastruct. Res.* **26**, 31-43 (1969).
22. Reynolds, E. S. *J. Cell Biol.* **17**, 208-212 (1963).

Two phospholipase pools for prostaglandin synthesis in macrophages

Wei Hsueh, Usha Desai, F. Gonzalez-Crussi, Ruth Lamb & Agnes Chu

The Department of Pathology of The Children's Memorial Hospital, Northwestern University Medical School, Chicago, Illinois 60614, USA
The Department of Pathology of University of Connecticut Health Center, Farmington, Connecticut 06032, USA

Macrophages in culture produce prostaglandins in response to a variety of phagocytic and non-phagocytic stimuli¹⁻⁸. As prostaglandins are not stored in cells, and mammalian cells contain very little free arachidonic acid, synthesis and release of prostaglandins depends on the release of the precursor, arachidonic acid, from cell lipids. Many agents that stimulate cell prostaglandin production act by releasing arachidonic acid⁹, presumably by activating phospholipase A₂ (refs 10, 11) or phospholipase C (refs 12, 13), depending on the cell system used. We have shown previously that rabbit alveolar macrophages secrete arachidonic acid as well as prostaglandins in response to phagocytic stimuli^{2,4,6}. This secretion depends not on particle attachment, but rather on interiorization of the particles⁴. Furthermore, the time course of prostaglandin and arachidonic acid secretion does not parallel that of particle engulfment *per se*, but of the release of lysosomal enzymes⁴, indicating that the release of arachidonate and prostaglandin is associated with the latter. We now describe experiments which suggest that these are two independent pools of phospholipases in macrophages, one in the lysosomes and one elsewhere.

To assess the prostaglandin release, we prelabelled the cell lipids of cultured rabbit alveolar macrophages (more than 90% macrophages) with ^{14}C -arachidonic acid (Amersham) as previously described². Analysis of the cell lipids showed a pattern of incorporation similar to that of our previous studies²: approximately 40–50% of the radioactive arachidonic acid was incorporated into cell lipids after 1 h of incubation. Most of the radioactivity (75%) was found in phospholipids, particularly in phosphatidylcholine ($48 \pm 1\%$, $n = 10$) and phosphatidylethanolamine ($19 \pm 1\%$). A small amount also resided in phosphatidylserine and phosphatidylinositol ($8 \pm 1\%$) and neutral lipids ($5 \pm 1\%$). Less than 2% of the radioactivity in cells remained in the free fatty acid fraction.

Stimulation of macrophages with various phagocytic and non-phagocytic stimuli resulted in prostaglandin release. However, only a small amount of arachidonic acid could be detected in the medium in the absence of serum. This is probably due to immediate reincorporation of the unreacted arachidonic acid into phospholipids. To avoid this problem, we used fatty acid-free bovine serum albumin to 'trap' the released fatty acid^{2,14}. Thus, the activity of phospholipase could be quantitatively assessed.

We have previously demonstrated a close association between the release of lysosomal enzyme and that of arachidonic acid and prostaglandin by macrophages, following zymosan phagocytosis^{2,4}. This suggests a cause-effect relationship between these two mechanisms. Prostaglandin synthesis, however, does not regulate lysosomal enzyme release, as inhibition of the former did not affect the latter^{2,15}. As alveolar macrophage lysosomal granules do contain phospholipase A₂ (ref. 16), it is possible that phagosome-lysosome fusion, on releasing acid hydrolases, activates lysosomal phospholipase, thus explaining the close

association between release of arachidonic acid and lysosomal enzyme. To investigate further whether activation of lysosomal enzyme *per se* is sufficient to elicit prostaglandin synthesis, we examined the effect of C5 fragment, a chemotactic agent known to elicit lysosomal enzyme release¹⁷, on arachidonic acid release and prostaglandin synthesis. The effect of C5 fragment was then compared with that of zymosan phagocytosis. As shown in Fig. 1b, both zymosan and C5 fragment resulted in significant increased release of arachidonic acid, up to 4–5 times the control. However, these two agents produced different effects on the coupling efficiency of phospholipase (arachidonic acid release) and cyclooxygenase (prostaglandin synthesis). Figure 1a shows that prostaglandin synthesis was increased to 2.5 times the control value following zymosan stimulation, but only to 1.1 times control value after C5 fragment treatment. Thus, phagocytosis of zymosan resulted in tight coupling between arachidonic acid release and prostaglandin synthesis (arachidonate/prostaglandin ratio = 3, Fig. 1d), whereas C5 fragment, with an arachidonate/prostaglandin ratio of 7.7 (Fig. 1d), produced a disproportionately small prostaglandin synthesis for a large arachidonic acid release. Lysosomal enzyme release remained comparable with both stimulants (approximately three times the control value, Fig. 1c).

Because it is generally accepted that phospholipase, not cyclooxygenase, is the rate-limiting enzyme for prostaglandin synthesis, this dissociation between arachidonic acid release and prostaglandin synthesis by C5 fragment can only be explained as

Fig. 1 Release of ^{14}C -prostaglandin, ^{14}C -arachidonic acid and acid phosphatase by rabbit alveolar macrophages in response to different stimuli. Rabbit alveolar macrophages were collected by lung lavage and plated in Dulbecco's modified Eagle's medium (DMEM) containing 15% fetal calf serum as previously described² (5×10^6 cells per ml per plate). After overnight incubation, the cells were changed to serum-free DMEM. ^{14}C -arachidonic acid (specific activity 55 mCi mmol^{-1} , Amersham) was added to the cultures as the sodium salt at a concentration of $1 \times 10^6 \text{ c.p.m. ml}^{-1}$. The cultures were incubated for 1 h to label the cell lipids. The cells were washed twice with DMEM and changed to DMEM containing 2 mg ml^{-1} fatty acid-free bovine serum albumin. Different stimuli with or without indomethacin were added to the cells and were incubated for 2 h. At the end of incubation, the medium was collected, an aliquot (1/10) was saved for acid phosphatase determination by fluorometric methods as described earlier. The remainder was acidified to pH 3.5, extracted with ethyl acetate and applied to silica gel G plates². The solvent system used for separation of prostaglandins and arachidonic acid was the organic phase from ethyl acetate/water/iso-octane/acetic acid, 110:100:50:20 (ref. 25). Zones corresponding to various prostaglandin standards and 12-HETE (prepared according to Goetzl *et al.*²⁶) were scraped and counted in a liquid scintillation counter. Total prostaglandin release (sum of 6-keto $\text{PGF}_{1\alpha} + \text{PGF}_{2\alpha} + \text{PGE}_2 + \text{PGD}_2$, c.p.m.) was calculated (a), the zone corresponding to arachidonic acid standard was also scraped and counted (b) and the acid phosphatase activity in the medium was assayed (c). All values are expressed as percentage of unstimulated control. The intracellular acid phosphatase content of the macrophages treated by various stimuli was determined, and found to be the same as that of unstimulated cells (after correction was made for released amount). d, Ratio of release of arachidonase and prostaglandin. The chemotactic fragment of human C5 (C5 fragment) was isolated from human serum and assayed by a modification of Boyden chamber assay as described previously²⁷. The chemotactic activity was expressed in units of ED_{50} . One ED_{50} of chemotaxis is defined as the amount of C5 fragment required for half-maximal chemotactic response in Boyden chamber assay. Z, zymosan (1 mg ml^{-1}); I, indomethacin ($5 \mu\text{g ml}^{-1}$); C5_{fr}, C5 fragment ($75 \text{ ED}_{50} \text{ ml}^{-1}$); PMA, 10^{-6} M . The numbers within the circles give the numbers of experiments. All values are mean $\pm$ s.e.m.

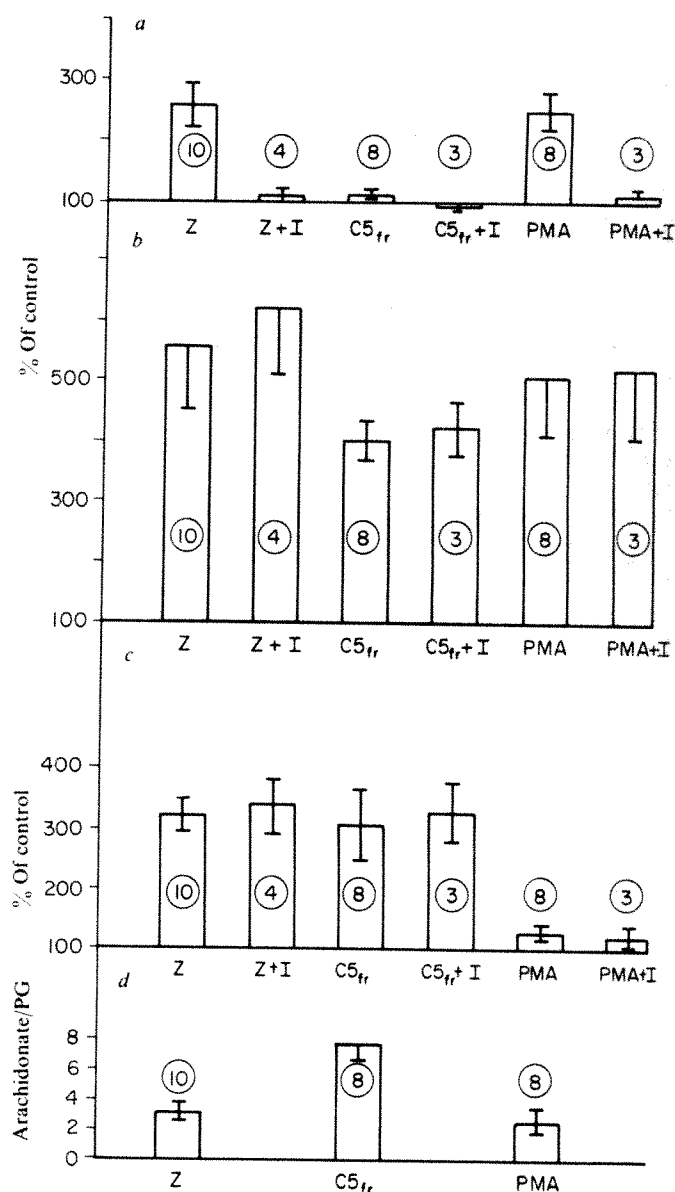


Table 1 Release of arachidonic acid and prostaglandins by alveolar macrophages in response to various stimuli

	Prostaglandins		Arachidonic acid	
	c.p.m.	% Of control	c.p.m.	% Of control
Unstimulated control (<i>n</i> = 9)	750 ± 93	100	898 ± 109	100
Zymosan (1 mg ml ⁻¹) (<i>n</i> = 9)	2,276 ± 283	301 ± 30	9,998 ± 1,326	994 ± 148
PMA (10 ⁻⁶ M) (<i>n</i> = 9)	2,079 ± 294	254 ± 28	7,542 ± 891	736 ± 55
C5 fragment (75 ED ₅₀) (<i>n</i> = 7)	1,008 ± 180	137 ± 21	6,613 ± 909	575 ± 62
Zymosan + PMA (<i>n</i> = 7)	2,942 ± 366	327 ± 45	24,466 ± 3,071	2,681 ± 371
Zymosan + C5 fragment (<i>n</i> = 4)	1,989 ± 265	278 ± 34	10,374 ± 2,059	967 ± 85
C5 fragment + PMA (<i>n</i> = 3)	2,271 ± 618	307 ± 25	38,489 ± 3,152	2,946 ± 384

Prostaglandin (PG) values are the sum of 6-keto PGF_{1α} + PGF_{2α} + PGE₂ + PGD₂(c.p.m.). For methods, see Fig. 1 legend. *n*, Number of experiments, each experiment in duplicate.

follows: (1) there is a switch of arachidonic acid metabolism to the lipoxygenase pathway, (2) C5 fragment inhibits cyclooxygenase, or (3) coupling of phospholipase and cyclooxygenase is prevented, as might occur when a physical distance is interposed between the two enzyme systems within the cell. We can easily rule out (1) as the production of 12-hydroxy-5,8,10,14-eicosatetraenoic acid (12-HETE) and 11-HETE (both comigrate in solvent system used in this study) by macrophages after C5 fragment stimulation is also very low (120 ± 9% of control compared with 247 ± 31% of control after zymosan stimulation, *n* = 5, mean ± s.e.m.). We cannot entirely rule out the second possibility, as C5 fragment slightly inhibited prostaglandin synthesis when added to zymosan-treated macrophages (Table 1). However, this inhibition is minimal, and cannot account for the discrepancy between arachidonic acid release and prostaglandin synthesis following C5 fragment stimulation. We consider the third explanation a more probable one. Cyclooxygenase activity usually resides in the microsomes, probably in the endoplasmic reticulum^{18,19}. Thus, activation of lysosomal phospholipase by C5 fragment resulted in release of arachidonic acid which remained inaccessible to cyclooxygenase. The phenomena implicated in phagocytosis are different. Zymosan engulfment leads to formation of phagosomes and many intracellular vacuoles²⁰, both of which contain plasma membrane and endoplasmic reticulum, on which cyclooxygenase presumably resides. Fusion of lysosomes with phagosomes and these vacuoles may be expected to result in approximation of activated phospholipase and cyclooxygenase. A consequence of this is more efficient coupling between these two enzyme systems.

Phorbol-myristate acetate (PMA), a membrane perturbant, stimulates the production of prostaglandins in MDCK cells²¹ and fibroblasts²². In rabbit alveolar macrophages, PMA, at the concentration of 10⁻⁶M (non-toxic to cells), resulted in high arachidonate and PG release (5- and 2.5-fold control value, respectively) and a very low lysosomal enzyme release (1.3 times control) (Fig. 1*a-d*), suggesting the existence of a phospholipase pool independent of lysosomes. Again, the large amount of prostaglandin synthesis is not due to decrease in lipoxygenase pathway, as the production of 11-HETE and 12-HETE is also very high (230 ± 16% of control, *n* = 5, mean ± s.e.m.). Morphological examination of these cells showed numerous

intracellular vacuoles produced by this chemical at a concentration range of 10⁻⁷–10⁻⁶M (Fig. 2*a*). Both prostaglandin release and vacuole formation became very reduced when the concentration of PMA was reduced to 10⁻⁸M (unpublished data). This observation suggests a close association between vacuole formation and prostaglandin production (Fig. 2*a*). We propose that: (1) the activated phospholipase resides outside the lysosomes, probably on the plasma membrane; (2) activation takes place during vacuole formation; and (3) the released arachidonic acid is immediately acted on by cyclooxygenase and lipoxygenase at the same site. Such a mechanism correlates with the observed efficient coupling between release of arachidonic acid and release of prostaglandin and HETEs. The small amount of lysosomal enzyme release could be due to fusion of lysosomes with the many cytoplasmic vacuoles. The association of vacuole formation and prostaglandin synthesis in mouse peritoneal macrophages stimulated by 12-O-tetradecanoylphorbol-13-acetate has been reported²⁰ as has the tight coupling of phospholipase and cyclooxygenase following stimulation by phorbol diester and other tumour promoters in cultured canine kidney cells^{23,24}. C5 fragment-treated cells are shown in Fig. 2*b* for comparison.

The existence of two pools of phospholipase is further supported by the additive effects of zymosan and PMA and of PMA and C5 fragment (Table 1). A combination of PMA (presumably activating membrane phospholipase) with either C5 fragment (presumably activating lysosomal phospholipase) or zymosan (presumably activating lysosomal phospholipase mainly) resulted in supramaximal increase of arachidonic acid release (26.8 times or 29.4 times the control). In contrast,

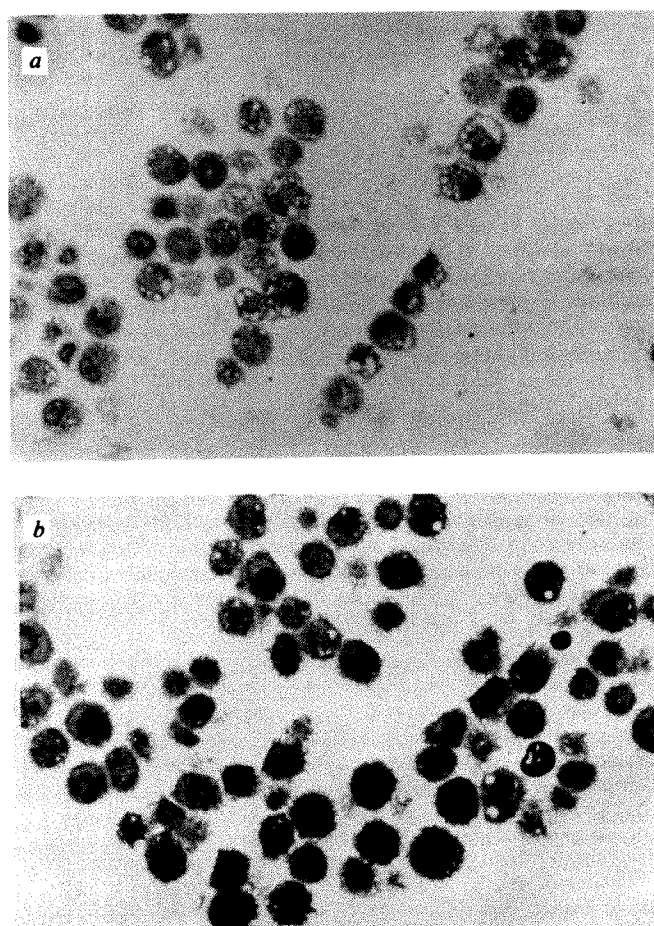


Fig. 2 PMA-treated macrophages (*a*) show numerous intracytoplasmic vacuoles. In contrast, C5 fragment-treated cells (*b*) show many cellular processes but less vacuole formation. The cells were fixed in 3% glutaraldehyde and embedded in epoxy resin. Sections (1 μm) were stained with toluidine blue.

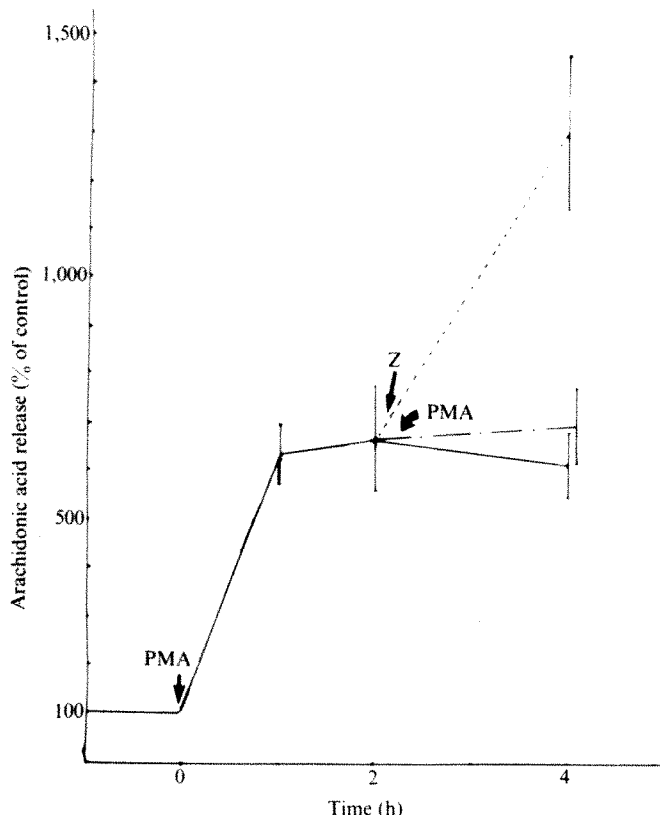


Fig. 3 Time course of arachidonic release by macrophages following various stimuli. At time zero, cells were stimulated with PMA (10^{-6} M). Solid line represents time course of arachidonic acid release following addition of PMA at time zero. Note flattening of the curve after 2 h of incubation. The shape of the curve was not altered despite a new dose of PMA (10^{-6} M) at 2 h (— · —). However, addition of zymosan (Z, 1 mg ml^{-1}) at this time caused a sharp increase of arachidonic acid release (— · — · —). All values are mean $\pm$ s.e.m., $n = 3$.

addition of zymosan with C5 fragment did not cause much more increase in arachidonic acid production than with zymosan alone. Note also that prostaglandin synthesis following combined zymosan and PMA treatment was only mildly elevated compared with either stimulus alone (Table 1). A possible explanation is saturation of cyclooxygenase, as both PMA and zymosan were used at near maximal dose in this experiment.

Our hypothesis may also be reconciled with the data illustrated in Fig. 3. Repeated stimulation of macrophages with PMA resulted in flattening of the arachidonate releasing curve, suggesting exhaustion of the non-lysosomal phospholipase pool. Addition of zymosan at this time elicited a second sharp increase of arachidonic acid release, indicating the activation of a different pool of phospholipases, namely the lysosomal phospholipases.

This work was supported by NIH grant 1 R23 AI 16147 and a grant from the American Lung Association. Dr John Pike (Upjohn) provided the prostaglandin standards.

Received 23 May 1980; accepted 28 January 1981.

- Humes, J. L. *et al.* *Nature* **269**, 149–151 (1977).
- Hseuh, W., Kuhn, C. III & Needleman, P. *Biochem. J.* **184**, 345–354 (1979).
- Brune, K., Glatt, M., Kälin, H. & Peskar, B. A. *Nature* **284**, 261–263 (1978).
- Hseuh, W., Gonzalez-Crussi, F. & Hanneman, E. *Nature* **283**, 80–82 (1980).
- Grimm, W., Seitz, M., Kirchner, H. & Gerns, D. *Cell. Immun.* **40**, 419–426 (1978).
- Hseuh, W. *Am. J. Path.* **96**, 137–147 (1979).
- Bonney, R. J., Naruns, P., Davies, P. & Humes, J. L. *Prostaglandins* **18**, 605–616 (1979).
- Passwell, J. H., Dayer, J. M. & Merler, E. *J. Immun.* **123**, 115–120 (1979).
- Levine, L., Pong, S. S., Hong, S. C. & Tam, S. in *Biochemical Aspects of Prostaglandins and Thromboxanes* (eds Kharasch, N. & Fried, J.) 15–38 (Academic, New York, 1977).
- Flower, R. J. & Blackwell, G. J. *Biochem. Pharmacol.* **25**, 285–291 (1976).
- Vogt, W. in *Advances in Prostaglandin and Thromboxane Research* Vol. 3 (eds Galli, C., Galli, G. & Porcellati, G.) 89–95 (Raven, New York, 1978).
- Kennerly, D. A., Sullivan, T. J., Sylvester, P. & Parker, C. W. *J. exp. Med.* **150**, 1039–1044 (1979).
- Bell, R. L., Kennerly, D. A., Stanford, N. & Majerus, P. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3238–3241 (1979).
- Isakson, P. C., Raz, A., Denney, S. E., Wyche, A. & Needleman, P. *Prostaglandins* **14**, 853–871 (1977).

- Bonney, R. J. *et al.* *Biochem. J.* **176**, 433–442 (1978).
- Franson, R. C. & Waite, M. J. *Cell Biol.* **56**, 621–627 (1973).
- McCarthy, K. & Henson, P. M. *J. Immun.* **123**, 2511–2517 (1979).
- Kataoka, K., Ramwell, P. W. & Jessup, S. *Science* **157**, 1187–1189 (1967).
- Bohman, S. O. & Larsson, C. *J. ultrastruct. Res.* **50**, 367 (1975).
- Brune, K., Glatt, M. & Kälin, H. *Nature* **274**, 261–263 (1978).
- Ohuchi, K. & Levine, L. *J. biol. Chem.* **253**, 4783–4790 (1978).
- Brinckerhoff, C. E., McMillan, R. M., Fahey, J. V. & Harris, E. D. *Arthritis Rheum.* **22**, 1109–1116 (1979).
- Levine, L. & Hassid, A. *Biochem. biophys. Res. Commun.* **79**, 477–484 (1977).
- Levine, L. & Ohuchi, K. *Nature* **276**, 274–275 (1978).
- Hamberg, M. & Samuelsson, B. *J. biol. Chem.* **241**, 257–263 (1966).
- Goetzl, E. J., Woods, J. M. & Gorman, R. R. *J. clin. Invest.* **59**, 179–183 (1977).
- Desai, U., Kruetzer, D. L., Showell, H., Arroyave, C. V. & Ward, P. *Am. J. Path.* **96**, 71–84 (1979).

Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies

Walter Gerhard, Jonathan Yewdell & Mark E. Frankel

The Wistar Institute of Anatomy and Biology, Philadelphia, Pennsylvania 19104, USA

Robert Webster

St Jude Children's Research Hospital, Memphis, Tennessee 38101, USA

The recurrence of influenza virus infection in man is attributed primarily to changes occurring in the antigenic structure of the viral surface glycoproteins, especially of the haemagglutinin (HA) molecule^{1–4}. Comparative antigenic analysis⁵ of epidemic influenza virus strains has allowed the description of 'strain-specific' and 'cross-reactive' antigenic determinants^{6–8}. However, the interpretation of these findings remained ambiguous, because the specificity of the applied antisera was insufficiently defined and because the antigenic differences among the HA molecules of various epidemic virus strains resulted presumably from a large number of amino acid substitutions^{9–11}. Thus, in characterizing the antigenic structure of the HA molecule, our approach has been (1) to generate a panel of monoclonal anti-HA hybridoma antibodies, (2) to use some of these antibodies to select mutants of the influenza A/PR/8/34 (PR8) virus expressing antigenically altered HA molecules, and (3) to construct an operational antigenic map of the HA molecule by comparative antigenic analysis of the mutant viruses with the monoclonal antibodies. As we report here, analysis of the 34 mutant viruses selected has enabled us to define four antigenic sites on the HA molecule. Our observation that these sites have undergone antigenic drift to a different extent in nature implies that the mechanisms responsible for antigenic drift act selectively on distinct structures of the HA molecule.

Fifty-eight anti-HA hybridoma antibodies were produced by fusion of lymphocytes from spleen and/or mediastinal lymph nodes of 12 BALB/c mice after intraperitoneal or intranasal inoculation with influenza virus¹². PR8 virus mutants with antigenic alterations (each resulting probably from a single point mutation¹³) in the HA molecule were then obtained by incubating cloned PR8 virus seed with selected hybridoma antibodies and by growing the non-neutralized virus fraction in embryonated hen eggs or in the allantois on shell culture system^{14,15}. Antigenic analysis performed by radioimmunoassay (RIA)¹⁶ on each set of mutant viruses (exemplified in Table 1 with mutants (CV) selected with antibody Sa-1) clearly showed that the non-neutralized viruses are antigenic mutants exhibiting an alteration in the HA molecule which reduced the capacity of the selecting antibody (in present example Sa-1) to bind to the mutant viruses in the RIA. Several antigenically distinct mutant viruses were usually selected by any given hybridoma antibody. For example, antibodies Sa-8 and Sa-9 show that mutants CV2 and CV1 are antigenically unique, whereas antibody Sa-14 shows that CV5 differs from CV6. As no monoclonal antibody

Table 1 Antigenic analysis of the mutant viruses selected in the presence of antibody Sa-1

Hybridoma antibody	Amount of antibody (log 2) binding in RIA to PR8 and its mutants				
	PR8-Mt.S. (parental)	CV1	CV2 (CV9)	CV5	CV6 (CV7, CV8, CV10)
Sa-1 (C)	=6.64	<1	<1	<1	<1
Sa-8	=6.64	6.46 ± 0.24	1	6.26 ± 0.45	6.58 ± 0.18
Sa-14	=6.64	<1	2.48 ± 0.98	<1	6.35 ± 0.02
Sa-9	=6.64	2.09 ± 0.94	6.58 ± 0.31	5.71 ± 0.25	5.78 ± 0.65
36 antibodies	=6.64	6.41 ± 0.24	6.36 ± 0.20	6.40 ± 0.13	6.54 ± 0.12

PR8 mutants were selected by titration of PR8-Mt.S. virus seed in the presence of anti-HA hybridoma ascitic fluid in the allantoic cavity of 10-day-old embryonated hen eggs¹⁴ or in the allantois on shell culture system¹⁵. Virus grown at the titration end point was collected and recloned twice in the absence of antibody in embryonated hen eggs. The cloned viruses were grown up to large quantity, purified and quantitated by HA titration. Viral immunoadsorbents were prepared as follows: 250 µl of HAS (phosphate-buffered saline, 0.04% sodium azide) containing 8×10^4 haemagglutinating units (HAU) of purified virus was mixed with 250 µl of HAS containing 2% NP40, incubated for 30 min at 37 °C and then diluted 1:10 with HAS. These viral NP40 stock solutions could be stored at 4 °C without detectably losing their antibody-binding capacity. The viral immunoadsorbents were prepared by incubating wells of polyvinyl microtitre plates overnight at room temperature with 25-µl aliquots of 1:100 dilutions of the viral NP40 stock solutions (corresponding to 4 HAU of undisrupted virus per well). The plates were incubated for 1 h with HAS containing 1% bovine serum albumin before use in RIA. Supernatants from hybridoma cultures were titrated in the RIA in serial 1:2 dilution steps against the parental viral immunoadsorbent. A standard binding curve was generated in this way for each antibody. One dilution of each antibody (upper part of approximately linear range of the standard curve) was then tested in the RIA against parental and the mutant viral immunoadsorbents. The c.p.m. observed in the parental interaction were defined as 100% (log 2 = 6.64). The antibody-mutant interactions were then expressed as a fraction (log 2) of the standard binding curve of the antibody-parental virus interaction. In each assay a given antibody was tested in parallel on the parental virus (quadruplicate samples) and on the individual mutants (duplicates) of a mutant set. Each assay was performed at least three times. The data in Table 1 are expressed as the mean ± standard deviation of three or four assays. The average standard deviation of antibody-mutant binding values (in the binding range of 4.64–6.64) is 0.37. This value was computed from 35 individual antibody-mutant interactions that were repeated independently of each other over ~1 yr.

was able to discriminate between either CV2 and CV9 or between CV6, CV7, CV8 and CV10, it is possible that the mutant viruses within each of these groups exhibit the same mutation. Lastly, the finding that 36 of the 58 anti-HA antibodies reacted equally well with parental virus and any of the C variants indicates that the antigenic changes induced by individual mutations represent minor modifications in a largely conserved antigenic background.

The observation that individual mutations concomitantly modify certain epitopes¹⁷ (structure on the antigen directly complementary to the combining site of a given antibody) provides an operational basis for the construction of an epitope map^{18,19}. Thus, in the example given above, a relationship between the epitopes delineated by antibodies Sa-1, Sa-8, Sa-9 and Sa-14 is clearly shown by the observation that these epitopes were changed by one or several of the mutations exhibited by the CV mutant viruses. This approach for mapping the antigenic structure of the HA molecule is limited, however, by the number and types of distinct antigenic mutants available for comparative analyses. Thus, the epitopes delineated by the 36 hybridoma antibodies which reacted equally well with the parental virus and the CV mutants could not be mapped on the basis of the CV-mutant virus analysis. We have proceeded, therefore, to select from the same cloned PR8 virus seed additional mutants in the presence of 10 other monoclonal anti-HA antibodies. Each set of mutants was then subjected to an antigenic analysis as exemplified in Table 1 with the CV mutants. A total of 75 PR8 mutants were analysed, of which 34 exhibited a unique antigenic modification. The other 41 mutants could not be further differentiated from a corresponding unique mutant by the present antibody panel and thus probably represent multiple isolates of those unique mutants.

The composite results of the comparative antigenic analyses of the mutants are shown schematically in Fig. 1 where each row represents the reactivity of an anti-HA antibody in RIA with the parental virus (column 1) and the 34 antigenically unique mutants (columns 2–35). Three antibody-binding values are distinguished, black, cross-hatched and white squares representing >50, 25–50 and <25% of the binding to parental virus, respectively. All the hybridoma antibodies, except antibodies Cx-1 to Cx-8, exhibit reduced binding to one or several of the virus mutants. Antibodies and mutant viruses could thus be assembled into four complementary antibody-virus mutant groups designated Sa, Sb, Ca and Cb (the choice for the designations will become clear in the next section). For instance, the mutant viruses CV, KV and PV were assembled into the mutant group Sa while complementary antibodies delineating epitopes that were predominantly modified by antigenic changes exhibited by this mutant group were assembled into a family of antibodies accordingly designated Sa-1 to Sa-15.

The resulting epitope map indicates that each mutant group characterizes an antigenic structure of the HA molecule which can undergo largely independent antigenic alterations. For example, mutations in the Sa mutant group affect the binding capacity of the Sa but not of the Ca antibody family and vice versa. Thus, adopting Urbanski and Margoliash's²⁰ definition of a discrete antigenic site as one to which the binding of specific antibodies is not affected by residue changes at the neighbouring sites, we tentatively designated these structures of the HA molecule as sites Sa, Sb, Ca and Cb. We stress that the reciprocal relationship between antibody families and virus mutant groups is only compatible with the notion that the mapping of antibodies is based on differences in antibody specificity and not on differences in avidity of antibodies directed against the same antigenic site. Also, the very low frequency of mutants selected in the presence of two antibodies directed against different sites further supported the notion of the independent mutability of these sites (unpublished observations).

Although most antibodies could be mapped to a single antigenic site, a few exhibited reduced binding to mutants assigned to different sites. In most cases, the observed overlap was between two sites (especially Sa and Sb) and, as discussed below, probably indicates the close physical proximity of these sites. In a few cases, however, linkages among three or four sites were observed. Notably, these linkages were only evident at the level

Table 2 The frequency of mutants in individual sites

Hybridoma antibody	No. of virus samples analysed	Frequency of mutants	
		In individual HA epitopes (log 10, mean ± s.e.m.)	Average per HA site
Sa-1 (C)	4	–5.59 ± 0.09	
Sa-4 (P)	1	–6.49	–6.14 ± 0.41
Sa-6 (K)	2	–6.08 ± 0.60	
Sa-15	3	–6.41 ± 0.33	
Sb-6 (E)	2	–5.32 ± 0.19	–5.48 ± 0.22
Sb-7 (B)	4	–5.63 ± 0.18	
Ca-1 (N)	5	–5.21 ± 0.34	
Ca-4 (D)	5	–4.86 ± 0.21	–5.31 ± 0.50
Ca-6 (S)	1	–5.85	
Cb-13 (L)	4	–6.26 ± 0.19	–5.98 ± 0.40
Cb-5 (R)	2	–5.70 ± 0.67	

Cloned PR8 virus was titrated in the allantois on shell culture system in the absence and presence of an 'over-neutralizing' dose of the indicated anti-HA antibodies¹⁵. Viruses that grew at the titration end point in the presence of anti-HA antibody were shown to be mutants by subsequent antigenic analysis. The mutant frequency (ratio of virus titre in the presence of anti-HA antibody to virus titre in its absence) was determined in several independently cloned PR8 virus samples. Letters in parentheses indicate the viral mutants selected (see Fig. 1).

these heterologous virus strains, suggesting a varying extent of antigenic drift in nature. The Sb antibody family is almost completely nonreactive with virus strains isolated after PR8. Thus, the antigenic site delineated by the Sb antibodies underwent drastic antigenic drift in nature while drift occurred to a slightly lesser extent in the Sa site, in which four cross-reacting antibodies react with heterologous viruses isolated shortly before (SW and WSN, in 1931 and 1933) or shortly after (MEL and BH, in 1935) the isolation of PR8. In contrast, drift occurred

in nature to a much lesser extent in the antigenic sites delineated by the Ca, Cb and Cx antibodies. (Based on these observations we have designated the PR8 'strain-specific' antigenic sites Sa and Sb and the 'cross-reactive' sites Ca and Cb. Note, however, that the Sa and Cb sites contain both 'specific' and 'cross-reactive' epitopes, although at different proportions.)

If antigenic drift in nature results from the selection of spontaneously arising mutant viruses that exhibit a favoured growth potential amid a pre-existing immunity in humans⁴, the mechanisms responsible for antigenic drift clearly act differently on the various antigenic sites defined here. This phenomenon is probably not due to an inherently greater mutability of the antigenic structures corresponding to the Sa and Sb sites as the average frequency of antigenic mutants present in cloned egg-grown PR8 virus is similar in 'specific' and 'cross-reactive' antigenic sites (Table 2). Therefore, the antiviral immune mechanisms seem to exert a stronger selection pressure for mutations in the Sa and Sb sites than the Ca and Cb sites. This could result from a predominant immune response of humans to the Sa and Sb sites and/or greater protective antiviral activities mediated by these antibodies.

Preliminary experiments showed that antibodies directed against the Sa and Sb sites exhibit *in vitro* a >10-fold higher virus-neutralizing potency than antibodies directed against the Cb site, which could explain the more extensive drift in the Sa and Sb sites than in the Cb site. However, very little drift occurred in the Ca site despite the relatively high virus-neutralizing potency of anti-Ca antibodies *in vitro*. Perhaps humans are unable to produce antibodies against the Ca site in sufficient quantities, or anti-Ca antibodies are ineffective *in vivo*.

We have thus used hybridoma antibodies and virus mutants selected *in vitro* to characterize the antigenic topology of the PR8 HA molecule. Webster and Laver¹⁹ have used a similar approach to demonstrate the existence of at least three non-overlapping antigenic sites on the HA molecule of influenza virus A/Mem/71. This approach has so far been used only to investigate the antigenic structure of the influenza virus HA molecule but it is clearly applicable to other viral proteins and to surface proteins of eukaryotic and prokaryotic organisms³⁰. Thus, it should be possible to determine the antigenic structure of a wide variety of biologically relevant antigens, which in turn may be useful for understanding the relationship between specificity and function of immune effector mechanisms.

We thank Maureen Carey, Dale McCreedy and Alex Taylor for help with these experiments. The study was supported in part by grants from the National Institute of Allergy and Infectious Diseases (AI 13989), the National Multiple Sclerosis Society (RG 851C6) and the Medical Scientist Training Program (GM 07170).

Received 10 November 1980; accepted 23 February 1981.

1. Dowdle, W. R., Coleman, M. T. & Gregg, M. B. *Prog. med. Virol.* **17**, 91–135 (1974).
2. Stuart-Harris, C. *Br. med. Bull.* **35**, 3–8 (1979).
3. Kilbourne, E. D. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 483–538 (Academic, New York, 1975).
4. Webster, R. G. & Laver, W. G. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 270–315 (Academic, New York, 1975).
5. Reichlin, M. *Adv. Immun.* **20**, 71–123 (1975).
6. Schild, G. C. *J. gen. Virol.* **9**, 191–200 (1970).
7. Laver, W. G., Downie, J. C. & Webster, R. G. *Virology* **59**, 230–244 (1974).
8. Virelizier, J. L., Allison, A. C. & Schild, G. C. *J. exp. Med.* **157**, 1571–1578 (1974).
9. Laver, W. G., Air, G. M., Doppeide, T. A. & Ward, C. W. *Nature* **283**, 454–457 (1980).
10. Verhoeven, M. *et al. Nature* **286**, 771–776 (1980).
11. Gething, M.-J., Bye, J., Skehel, J. & Waterfield, M. *Nature* **287**, 302–306 (1980).
12. Gerhard, W., Yewdell, J. W., Frankel, M. E., Lopes, A. D. & Staudt, L. in *Monoclonal Antibodies* (eds Kennett, R. H., McKearn, T. J. & Bechtol, K. B.) 317–333 (Plenum, New York).
13. Laver, W. G. *et al. Virology* **98**, 226–237 (1979); *Proc. natn. Acad. Sci. U.S.A.* **76**, 1425–1429 (1979).
14. Gerhard, W. & Webster, R. G. *J. exp. Med.* **148**, 383–392 (1978).
15. Yewdell, J. W., Webster, R. G. & Gerhard, W. *Nature* **279**, 246–248 (1979).
16. Frankel, M. & Gerhard, W. *Molec. Immun.* **16**, 101–106 (1979).
17. Jerne, N. K. *Annls Immun. Inst. Pasteur, Paris* **125**, 373–389 (1974).
18. Gerhard, W., Yewdell, J. W. & Frankel, M. E. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 273–281 (Elsevier, Amsterdam, 1980).
19. Webster, R. G. & Laver, W. G. *Virology* **104**, 139–148 (1980).
20. Urbanski, G. J. & Margoliash, E. *J. Immun.* **118**, 1170–1180 (1977).
21. Lubeck, M. & Gerhard, W. *Virology* (in the press).
22. Eckert, E. A. *J. Virol.* **11**, 183–192 (1973).
23. Brand, C. M. & Skehel, J. J. *Nature* **238**, 145–147 (1972).

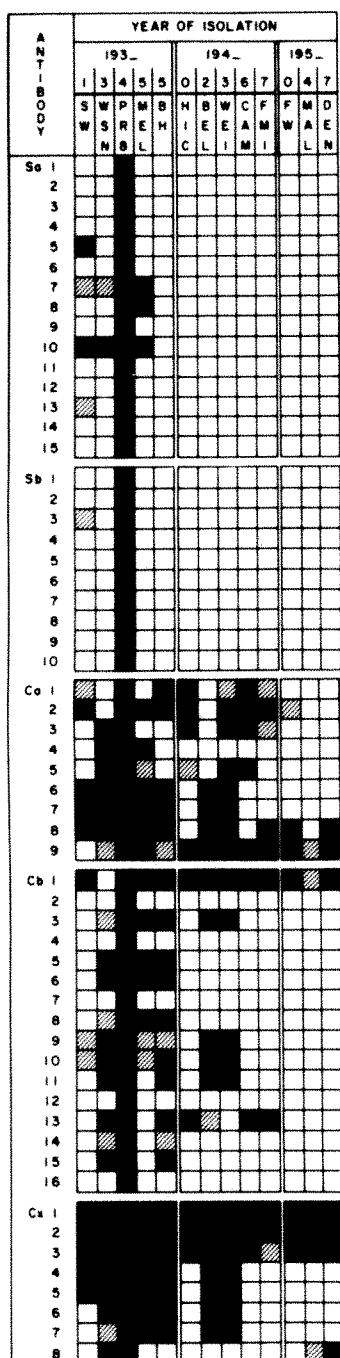


Fig. 2 Binding of hybridoma antibodies in RIA to epidemic influenza virus isolates. The data are similar to those of Fig. 1 except that the white squares represent binding of <2.64, the cross-hatched squares 2.64–4.64, and the black squares >4.64. Wider binding ranges than those in Fig. 1 were chosen so as to differentiate between low (2.64–4.64) and non-detectable (<2.64) cross-reaction. The virus strains used as immunoadsorbents are arranged according to the year of their isolation: A/SW/31(H1N1); A/WSN/33(H1N1); A/PR/8/34(H1N1); A/Melbourne/35(H1N1); A/BH/35(H1N1); A/Hickox/40(H1N1); A/Bellamy/42(H1N1); A/Weiss/45(H1N1); A/CAM/46(H1N1); A/FM/1/47(H1N1); A/FW/1/50(H1N1); A/Malaya/302/54(H1N1); A/Denver/1/57(H1N1).

24. Jackson, D. C., Dopheide, T. A., Russell, R. J., White, D. O. & Ward, C. W. *Virology* **93**, 458–465 (1979).
25. Jackson, D. C., Brown, L. A., White, D. O., Dopheide, T. A. & Ward, C. W. *J. Immun.* **123**, 2610–2617 (1979).
26. Laver, W. G. *et al.* in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 295–306 (Elsevier, Amsterdam, 1980).
27. Moss, B. A., Underwood, P. A., Bender, V. J. & Whittaker, R. G. in *Structure and Variation of Influenza Virus* (eds Laver, W. G. & Air, G. M.) 329–338 (Elsevier, Amsterdam, 1980).
28. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366–373 (1981).
29. Wiley, D. C., Wilson, I. A. & Skehel, J. J. *Nature* **289**, 373–378 (1981).
30. Holtkamp, B., Cramer, M., Lemke, H. & Rajewsky, K. *Nature* **289**, 66–68 (1981).
31. Yewdell, J. W. & Gerhard, W. A. *Rev. Microbiol.* (in the press).

Short-lived cytoplasmic regulators of gene expression in cell cybrids

C. Ronald Kahn*, Roger Bertolotti, Maryvonne Ninio & Mary C. Weiss

Centre de Genetique Moleculaire du CNRS,
91190 Gif-sur-Yvette, France

Somatic cell hybridization is a valuable tool for investigating the control of gene expression in eukaryotic cells^{1–4}. Studies of hybrid cells, heterokaryons, reconstructed cells and cybrids (cytoplasmic hybrids) have suggested that cytoplasmic factors may be involved in this regulatory process. Unfortunately, studies of this kind usually require that hybrid or modified cells be maintained for some time in a selective environment during which chromosomal losses or other changes may modify the genetic functions of the cells and thus vitiate conclusions about the mechanism of gene regulation. We report here the preparation of cybrids between enucleated mouse fibroblasts (Cl-1-D) and differentiated rat hepatoma cells (Fao) and the use of a combination of histological techniques to identify these modified cells early after fusion without the use of selective media. We found that albumin production in most cybrids was suppressed (extinguished) at 12–20 h after fusion but was restored by 48 h. These results suggest that there is a cytoplasmic factor in the fibroblast which exerts negative control over expression of the albumin gene, but which in the absence of the fibroblast nucleus, is not renewed and therefore short-lived.

The loss of gene functions characteristic of differentiation is commonly observed after cell fusion and is termed extinction^{1,2,4–6}. For example, most hybrids between fibroblasts and hepatoma cells do not express differentiated liver functions such as albumin synthesis, tyrosine aminotransferase and aldolase. The phenomenon, well studied in hybrids, has also been observed in heterokaryons (refs 7–10 and M. Ninio and M. Weiss, in preparation), and it has been suggested that differentiated gene functions are controlled by diffusible substances whose effects are negative or suppressive^{1–3}. Evidence that the substances concerned may be cytoplasmic has been provided by experiments¹¹ in which cybrids obtained by fusing Friend erythroleukaemia cells with cytoplasm of neuroblastoma cells or fibroblasts failed to produce haemoglobin in circumstances in which erythroleukaemic cells would have done so. However, most attempts to transfer extinction or other properties under genetic control by hybridization have failed (refs 12–14 and C. R. Kahn, T. V. Gopalakrishnan and M. C. Weiss, in preparation). Such negative experiments, however, do not rule out the possibility of cytoplasmic regulators of gene expression, because the cybrids usually cannot be examined until several days to weeks after fusion when grown in selective conditions. Here we used a combination of histological techniques to search for cytoplasmic regulatory factors early after cybrid formation.

The cytoplasm of the mouse fibroblast Cl-1-D was labelled by allowing the cells to phagocytose polystyrene beads (0.5–1.0 µm), after which enucleation was performed using the method of Gopalakrishnan and Thompson¹⁵. Cl-1-D cells are somewhat resistant to cytochalasin-induced enucleation, but with this method between 50 and 70% of the cells can be enucleated with high yield. After a short period of recovery, the cytoplasts, still attached to the plastic disks used for enucleation, were fused with the differentiated rat hepatoma, Fao, using polyethylene glycol (PEG). At selected time intervals, the cells were fixed and stained by an indirect immunofluorescent technique for albumin⁹ followed by nuclear staining with the fluorochrome dye Hoescht 33258 (ref. 16). The cells were then examined using phase contrast and epifluorescent illumination.

This combination of histological techniques allowed the identification of both parent cell types, heterokaryons, hybrids and cybrids. The fibroblast cytoplasm was labelled with beads and the hepatoma was not; the rat (hepatoma) nuclei stained homogeneously with Hoescht 33258, whereas mouse (fibroblast) nuclei possessed densely staining heterochromatin or chromocentres which fluoresced as bright spots (Fig. 1, top and middle rows respectively). Heterokaryons possessed cytoplasmic beads and both rat and mouse nuclei; hybrids, cytoplasmic beads and a large 'mixed-type' of nucleus; and cybrids, cytoplasmic beads and only a rat nucleus (or nuclei) (Fig. 1, bottom row). Each of these cell types was evaluated for albumin production by immunofluorescent staining (Fig. 1c, f, i).

At 12–20 h after fusion, cybrids between the enucleated fibroblast Cl-1-D and the rat hepatoma Fao showed a marked extinction of albumin production based on immunofluorescent staining (Fig. 1i). In four separate experiments, 59–72% of all mononuclear (1N) cybrids examined were negative for albumin, compared with <10% negative in normal cultures of Fao (Table 1; Fig. 2). This is comparable with the frequency of extinction of albumin production observed in the same experiments in heterokaryons between intact Cl-1-D and Fao at this time point (data not shown). The extinction was not due to a toxic effect of the PEG or to the fusion itself because intact Fao cells on the same disk or Fao homokaryons (Fao cells fused with other Fao) were between 89 and 91% positive for albumin by this technique. Furthermore, the extinction was not due to quenching of fluorescence by the polystyrene beads, because Fao cells which had been allowed to phagocytose beads directly showed comparable levels of albumin by immunofluorescence (data not shown).

Extinction was a rapid but transient phenomenon (Fig. 2, Table 1); within 6 h after fusion, 18–35% of all cybrids already demonstrated extinction, and extinction reached its maximum by 12–20 h (Table 1). After 24 h, a decline in the level of extinction was apparent, and by 45–48 h cybrids were 90% positive for albumin, a level comparable with that in the non-fused Fao.

Although we cannot accurately quantitate the ratio of cytoplasm contributed by the enucleated fibroblasts and the intact hepatoma, some attempt was made to evaluate the concentration dependence of the phenomenon by comparing the frequency of extinction in cybrids containing one nucleus with those containing more than one nucleus (but only rat hepatoma nuclei). In each experiment, and at every time point, the frequency of extinction was greater in the one-nucleus cybrids (1N), than in cybrids containing more than one nucleus (>1N) (and thus more than one rat hepatoma cell) (Table 1, Fig. 2b). In general, extinction seemed to be more frequent in those cybrids containing the largest numbers of beads, although this proved impossible to quantitate because cytoplasts varied considerably in the extent of labelling and those more heavily labelled with beads were preferentially lost during the enucleation procedure. However, these experiments strongly suggest that the cytoplasm of the mouse fibroblast contains a diffusible regulatory molecule which is short-lived and can influence the expression of albumin production in a time- and concentration-dependent fashion.

* Permanent address: Section on Cellular and Molecular Physiology, Building 10, Room 8S-243, National Institutes of Health, Bethesda, Maryland 20205, USA.

Despite intensive study, the molecular basis for the regulation of gene expression during the process of cellular differentiation remains poorly understood. It has been assumed that there may be both positive and negative regulators of gene expression in eukaryotic cells analogous to those observed in bacteria¹⁷. The existence of cytoplasmic molecules which may serve as regulators of gene expression has been clearly demonstrated by nuclear transplantation in protozoa and lower vertebrate species^{18,19}.

More recently, evidence has also been provided for the existence of similar positive and negative cytoplasmic regulatory factors in mammalian cells. Analysis of differentiated functions in somatic cell hybrids¹⁻⁶ and heterokaryons (refs 7-10 and M. Ninio and M. Weiss, in preparation) has suggested that the expression of genes in mammalian cells may be regulated by diffusible substances whose final effect may be either negative (extinction) or positive (activation). Evidence that these factors exist in the cytoplasm has been suggested by two types of experiment. Lipisch *et al.*²⁰ observed activation of the gene for tyrosine aminotransferase in reconstructed cells prepared from

mouse fibroblast nuclei and rat hepatoma cytoplasts, and Gopalakrishnan *et al.*^{11,21} have observed both extinction and activation of differentiated functions in cybrids prepared from mouse erythroleukaemia cells and cytoplasts derived from a variety of cells. With these exceptions, however, most attempts to transfer cytoplasmic regulatory factors have yielded negative results¹²⁻¹⁴. For example, we found that 6 weeks after fusion, cell hybrids between hepatoma and fibroblasts almost invariably showed extinction of differentiated liver functions (such as albumin and α -fetoprotein production, aldolase activity), whereas cybrids prepared using only the fibroblast cytoplasm did not, although chloramphenicol resistance (a fibroblast mitochondrial marker) was transferred, indicating that fusion had indeed taken place (C. R. Kahn, T. V. Gopalakrishnan and M. C. Weiss, in preparation).

Our data suggest that the fibroblast cytoplasm does indeed contain a molecule(s) capable of producing extinction of albumin production with very high efficiency. The reason this potential regulator of gene expression was not previously observed is that, in the absence of a nucleus coding for new synthesis, the

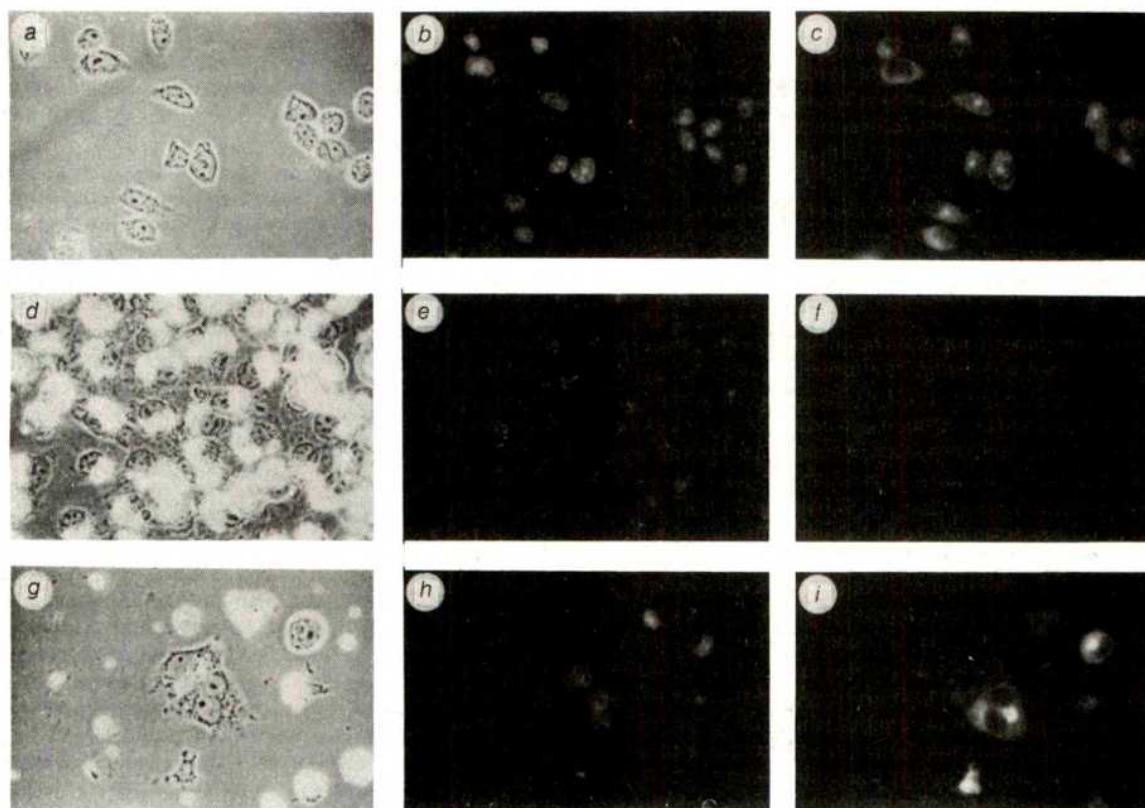


Fig. 1 Morphological and immunological characterization of the parental cells and cybrids. Cl-1-D (middle row) is a clonal line of fibroblastic mouse cells, sometimes referred to as LM(TK)²⁸. Fao (top row) is a fourth-generation clonal line of well differentiated rat hepatoma cells derived from the hepatoma line H4IIEC3 which is resistant to 8-azaguanine and ouabain²⁹. Both cell lines were cultivated in modified Ham's F12 medium³⁰ supplemented with 5% fetal calf serum in a humidified atmosphere, 7% CO₂. To prepare cybrids (bottom row), Cl-1-D cells were first allowed to phagocytose polystyrene beads (Dow) by adding a sterile suspension to culture medium and incubating the cells overnight. Beads of 0.8 μ m were of optimum size for easy detection without gross distortion of morphology *d*, although 0.5 and 1.0 μ m beads were also usable. The cells were then detached with trypsin and EDTA²⁹, resuspended in phosphate-buffered saline (PBS), and 1 ml of cell suspension ($2-3 \times 10^6$ cells ml⁻¹) was allowed to attach to 45-mm plastic disks precoated with concanavalin A (0.15 mg) and enucleated as described by Gopalakrishnan and Thompson¹⁵ using medium containing 10 μ g ml⁻¹ of cytochalasin B. A 50-70% enucleation could be achieved by centrifugation for 40 min at 16,000g at 37 °C with a ~40-50% yield. After a recovery period of 30-45 min in normal growth medium, the medium was aspirated and 1 ml of a suspension of $2.5-3.5 \times 10^5$ Fao cells was carefully layered onto the plastic disk. After 2-2.5 h at 37 °C in the incubator, the Fao cells were well attached and fusion was carried out by treating for 1 min with a solution of 50% PEG (1000 Merck, Ultrapure) in serum-free medium. The cells were washed four times with serum-free media, then returned to the incubator in 60-mm Petri dishes containing 5 ml of growth medium. At appropriate intervals, the coverslips were removed and prepared for histological examination as described by Ninio *et al.* (in preparation). The cells were fixed in 3% formaldehyde for 1 min, which was diluted 1:1 with methanol, then replaced completely by pure methanol for 30 min. The cells were then incubated with a rabbit anti-rat albumin serum (dilution 1:100) for 30 min at 37 °C, washed with PBS for 5 min, then stained with a fluorescein-labelled sheep anti-rabbit IgG (Institut Pasteur, dilution 1:200) for 30 min at 37 °C. After an additional wash with PBS, the cells were stained with the fluorochrome Hoechst 33258 (0.5 μ g ml⁻¹)¹⁶ and rinsed again with PBS. Glass coverslips were mounted over the cells with buffered glycerol and sealed with Histolique. The slides could then be stored at 4 °C or examined directly in phase contrast and epifluorescent illumination using a Zeiss UV microscope equipped with a HBO 200 W mercury lamp. Photographs were taken with either a $\times 40$ or $\times 100$ oil immersion objective, using Ilford HP-5 or Ektachrome film. This figure shows the appearance in phase contrast and fluorescence microscopy of the two parent cell types and the cybrids: *a, d, g*, phase contrast; *b, e, h*, Hoechst 33258; *c, f, i*, albumin. Hoechst 33258 stains the nuclei of the rat hepatoma homogeneously (*b*), whereas the mouse fibroblast nuclei have a dotted appearance due to more intense staining of the chromocentres (*e*). The bottom row shows a typical cybrid next to a binucleated Fao cell or homokaryon. Both cells have only rat type nuclei (*h*), but the cybrid also shows cytoplasmic beads as a result of fusion with a Cl-1-D cytoplast. Note that albumin (*i*) is extinguished in the cybrid but not in the adjacent Fao cell.

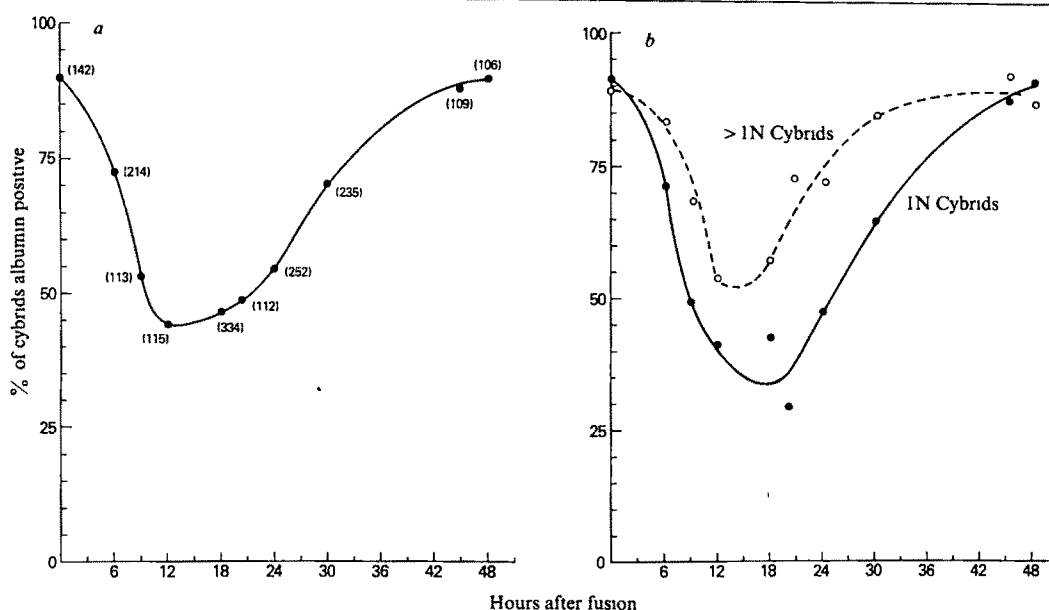


Fig. 2 Time course of extinction of albumin production in cybrids between Cl-1-D fibroblasts and the rat hepatoma Fao. *a*, The experiments in Table 1 have been summarized by totalling all cybrids examined at each time point for the four experiments and calculating the per cent albumin-positive cells by immunofluorescence. Numbers in parentheses indicate the number of cybrids examined. *b*, The data in *a* were subdivided to show the mononuclear cybrids (1N, solid line) versus the cybrids with more than one rat nucleus (>1N, broken line). Note that at every time point extinction is more frequent in the 1N cybrids. We believe that immunofluorescent staining for albumin will be positive until the cellular content of albumin is reduced by >75%.

factor has a relatively short half life. Thus, albumin production was inhibited or markedly repressed in the majority of cybrids by 12–20 h after fusion, but by 48 h extinction was no longer observed, presumably because the factor had been metabolized or diluted by cell division to a level no longer sufficient for activity. This is supported by data obtained in a parallel series of experiments in which heterokaryons between Cl-1-D and Fao cells showed extinction with a similar rate of onset, but which persisted for the 10 days of the study (M. Ninio and M. C. Weiss, in preparation). Although quantitation is difficult, both studies showed a dose-dependent effect; cybrids and heterokaryons with more than one hepatoma nucleus showed less extinction than those with only a single hepatoma nucleus.

Neither the chemical nature of the factor(s) nor the location of regulation is apparent. Previous studies with hybrids and mutant cells have suggested that regulation of gene expression occurs at a pretranslational level and is specific for the differentiated functions of the cell^{1–6}. Studies in prokaryotes have indicated that both proteins and small molecules, such as cyclic AMP, may act as regulatory factors¹⁷. The time course of extinction observed in these experiments is more consistent with the half life

of a protein than a cyclic nucleotide. However, it would not be surprising if the time course of effect varied extensively depending on the half life of the regulator in its new environment, as well as its mode of action. Activation of the tyrosine aminotransferase gene in the reconstructed cells formed from fibroblast nuclei and hepatoma cytoplasm lasted 6–8 weeks²⁰, and permanent activation of the gene for phenylalanine hydroxylase seems to occur in cybrids formed from erythroleukaemia cells and hepatoma cytoplasm²¹. However, using an experimental design similar to that reported here, except that cybrids were formed between enucleated hepatoma cells and fibroblasts, we have been unable to obtain evidence for any positive regulators of albumin gene expression.

Clearly, some cytoplasmic regulators of gene expression exist. The transient effect observed in the absence of the nucleus supports the concept that, in general, the epigenotype is stable¹. Thus, somatic-cell hybrids re-express parental differentiated functions following the loss of chromosomes from the extinguishing parent^{22–25}. It is also evident, however, that, in the eukaryotic cell, there are multiple potential mechanisms of gene regulation. Direct studies of cell extracts using chemical

Table 1 Immunofluorescent quantitation of albumin production in cybrids between Cl-1-D and Fao cells

Expt	Time (h)	No. of 1N cybrids examined		% Of 1N cybrids albumin positive	No. of >1N cybrids examined		% Of >1N cybrids albumin positive	No. of Total cybrids examined		% Of all cybrids albumin positive
		Alb ⁺	Alb ⁻		Alb ⁺	Alb ⁻		Alb ⁺	Alb ⁻	
1	6	64	38	62.3	9	2	81.8	73	40	64.6
	18	38	55	40.9	10	3	76.9	48	58	45.3
	30	58	19	75.3	29	2	93.5	87	19	82.1
2	18	20	43	31.7	18	26	40.9	38	69	35.5
	24	28	34	53.8	24	15	61.5	52	49	51.5
3	9	45	46	49.5	15	7	68.2	60	53	53.1
	20	17	43	28.3	38	14	73.0	55	57	49.1
	45	84	12	87.5	12	1	92.3	96	13	88.1
4	6	67	15	81.7	16	3	84.2	83	18	82.2
	12	31	46	40.3	15	13	53.6	46	59	43.8
	18	53	48	52.0	16	4	75.0	69	52	62.7
	24	53	53	50.0	22	3	88.0	75	56	67.6
	30	57	43	51.8	22	7	75.9	79	50	66.4
	48	76	8	91.6	19	3	86.4	95	11	89.6

fractionation and microinjection²⁶, as well as further characterization of the DNA and specific genes²⁷ from expressing and non-expressing cell lines, should elucidate these problems

Received 17 October 1980, accepted 15 January 1981

- 1 Ephrussi, B. *Hybridization of Somatic Cells* (Princeton University Press, New Jersey, 1972)
- 2 Weiss, M. C. *Proc 5th Int Congr hum Genet*, Mexico City, 284-292 (Excerpta Medica, Amsterdam, 1976)
- 3 Ringertz, N. R. & Savage, R. E. *Cell Hybrids* (Academic, New York, 1976)
- 4 Davidson, R. L. A. *Rev Genet* 8, 195-218 (1974)
- 5 Weiss, M. C., Sparkes, R. & Bertolotti, R. *Somat Cell Genet* 1, 27-40 (1975)
- 6 Deisseroth, A. *et al Somat Cell Genet* 2, 373-384 (1976)
- 7 Gordon, S. & Cohn, Z. *J exp Med* 131, 981-1003 (1970)
- 8 Thompson, E. B. & Gelehrter, T. D. *Proc natn Acad Sci U.S.A.* 68, 2589-2593 (1971)
- 9 Harris, H. *Cell Fusion The Dunham Lectures* (Clarendon, Oxford, 1970)
- 10 Zeuthen, J. S., Stenman, S., Fabricius, H. A. & Nilsson, K. *Cell Differentiation* 4, 369-386 (1976)
- 11 Gopalakrishnan, T. V., Thompson, E. B. & Anderson, W. F. *Proc natn Acad Sci U.S.A.* 74, 1642-1646 (1977)
- 12 Howell, A. N. & Sager, R. *Proc natn Acad Sci U.S.A.* 75, 2358-2362 (1978)
- 13 McBurney, M. W. & Strutt, B. *Expt Cell Res* 124, 171-180 (1979)
- 14 Wright, W. E. & Hayflick, L. *Expt Cell Res* 96, 113-121 (1975)
- 15 Gopalakrishnan, T. V. & Thompson, E. B. *Expt Cell Res* 96, 435-439 (1975)
- 16 Wright, W. E. & Gros, F. *Expt Cell Res* 111, 451-454 (1978)
- 17 Goldberger, R. F., Deeley, R. G. & Mullinix, K. P. *Adv Genet* 18, 1-68 (1976)
- 18 Danielli, J. F. *Harvey Lect* 58, 217-232 (1963)
- 19 DeRobertis, E. M. & Gurdon, J. B. *Proc natn Acad Sci U.S.A.* 74, 2470-2474 (1977)
- 20 Lipsch, L. A., Kates, J. R. & Lucas, J. J. *Nature* 281, 74-76 (1979)
- 21 Gopalakrishnan, T. V. & Anderson, W. F. *Proc natn Acad Sci U.S.A.* 76, 3932-3936 (1979)
- 22 Bertolotti, R. *Somat Cell Genet* 6, 579-602 (1977)
- 23 Klebe, R. J., Chen, T. R. & Ruddle, F. H. *Proc natn Acad Sci U.S.A.* 66, 1220-1227 (1970)
- 24 Weiss, M. C. & Chaplain, M. *Proc natn Acad Sci U.S.A.* 68, 3026-303C (1971)
- 25 Croce, C. M., L. twalk, G. & Koprowski, H. *Proc natn Acad Sci U.S.A.* 70, 1268-1272 (1973)
- 26 Diacumakos, F. G. *Meth Cell Biol* 7, 288-312 (1973)
- 27 Sperling, L. & Weiss, M. C. *Proc natn Acad Sci U.S.A.* 77, 3412-3416 (1970)
- 28 Dubbs, D. R. & Kit, S. *Expt Cell Res* 33, 19-28 (1964)
- 29 Deschatrette, J., Moore, E. E., Dubois, M., Cassio, D. & Weiss, M. C. *Somat Cell Genet* 5, 697-718 (1979)
- 30 Coon, H. G. & Weiss, M. C. *Proc natn Acad Sci U.S.A.* 62, 85-89 (1969)

Polyoma DNA sequences involved in control of viral gene expression in murine embryonal carcinoma cells

Michaël Katinka*, Marc Vasseur†, Nicole Montreau†, Moshe Yaniv* & Daniel Blangy†

* Département de Biologie moléculaire, Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France

† Laboratoire des Virus oncogènes, Université Pierre et Marie Curie, Institut de Recherches Scientifiques sur le Cancer, BP no 8, 94800 Villejuif, France

Polyoma virus can develop lytically in differentiated mouse cells, but is unable to express either early or late functions in undifferentiated early embryonic cells¹ or in embryonal carcinoma (EC) cells of the mouse. However, the EC cells become susceptible during differentiation and then behave as do mouse somatic cells². We have recently described the isolation of a new class of polyoma virus mutant (Py EC) able to develop normally in an EC cell line (PCC4-Aza) that is capable of trigeminal differentiation *in vivo*³ but exhibits a very limited differentiation *in vitro*⁴. The mutants do not grow in another EC cell line (F9) that is 'blocked'—'nullipotent' *in vivo*, but able to differentiate *in vitro* when grown in the presence of chemical inducers. Compared with their parental strain, all Py EC genomes analysed so far have an extra 20-50 bases in their *HpaII*-3 restriction fragment⁵. We now describe polyoma mutants that will grow in EC F9 cells, and the sequence changes in two of them. The region involved in the mutations can be folded into a putative tRNA-like secondary structure which may be involved in the regulation of viral early transcription with respect to the state of differentiation of the host cell.

Mutants of polyoma virus able to develop on F9 EC cells were obtained using two different procedures. F9 cells were infected (1) with an heterogeneous and highly defective polyoma stock, obtained by viral propagation at high multiplicity of infection (MOI) on permissive mouse cells, as already described for the

selection of Py EC on PCC4-Aza, or (2) with purified wild-type polyoma A2, at a MOI of 5,000 plaque-forming units (PFU) per cell. Following their infection, F9 cells were cultured as usual and scored periodically by immunofluorescence for polyoma T (tumour) and V (late structural proteins) antigen expression. In both procedures, and as reported for the PCC4-Aza cells, viral proteins were detected in 5-20% of the cells 3-4 weeks after infection. The virus produced was recovered and plaque purified three times on secondary mouse embryo cells (We propose calling the mutants Py EC, followed by the name of the cells on which they were selected and the isolate number, for example, Py EC PCC4-97, Py EC F9-5). Transforming ability on rat FR3T3 cells⁶, efficiency of tumour induction in newborn hamsters and burst size on various permissive mouse cell lines were checked for independent Py EC F9 isolates. As observed in the case of Py EC PCC4, no significant difference between Py EC F9 mutants and wild-type virus could be found.

Endonuclease restriction analysis of five independent Py EC F9 mutants revealed no common detectable size modifications of the *HpaII*-3 fragment (compare the EC PCC4 mutants of ref. 9) nor any specific rearrangement associated with either of the selection procedures used.

F9 EC cells infected with 10 PFU of any of the Py EC F9 mutants tested, exhibit, 42 h after infection, more than 60% of T-antigen positive cells, as measured by immunofluorescence⁸. In the same conditions of infection, never more than 5% PCC4-Aza cells were shown to express T or V antigen (results not shown).

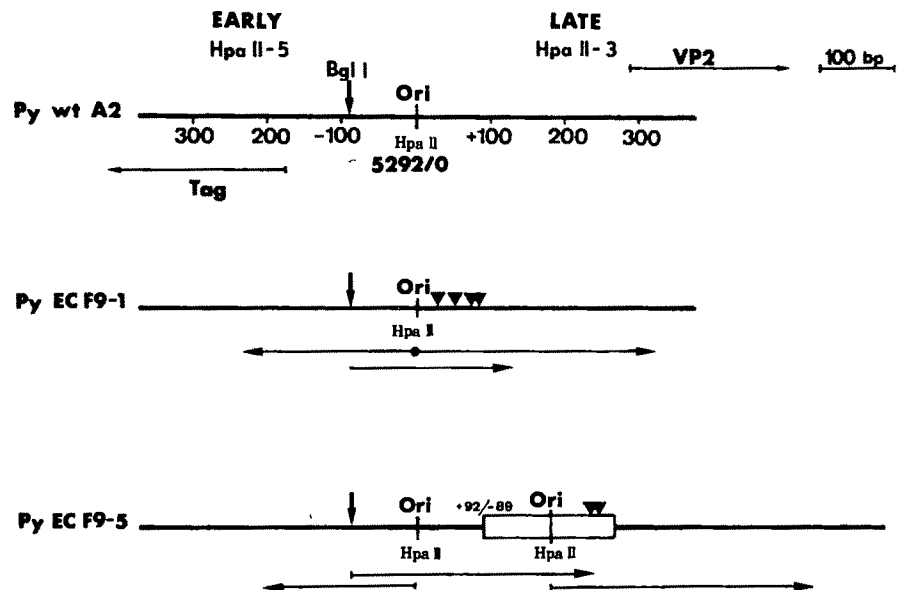
The efficiency of F9 infection by the Py EC F9 mutant can be compared with that of PCC4-Aza infection by Py EC PCC4 mutants⁹ in the latter case, at 37 °C, never more than 40% T- or V-antigen positive cells were observed (after infection at 100 PFU per cell) whereas F9 cells infected at 5-10 PFU with Py EC F9 mutants allow viral expression in 60-80% of the cells. All these biological properties suggest that restriction of polyoma expression in F9 and PCC4 may involve different factors.

The cloned viral genomes inserted into the *Bam*HI site of PBR322 were tested for their infectivity by transfecting secondary mouse embryo cells with viral DNA obtained after cleavage with *Bam*HI and ligation. The virus burst obtained was then used to infect both EC F9 and EC PCC4 cells at a MOI of 10-20 PFU per cell. When monitored for the appearance of T and V antigens by immunofluorescence, 60-80% of the EC F9 cells but only 3-5% of the EC PCC4 cells were positive. DNA extracted from two mutant infectious viruses revealed the expected restriction pattern, without apparent defective genomes.

About 500 nucleotides around the origin of DNA replication were sequenced in Py EC F9-1 and about 700 in Py EC F9-5, as described in the legends to Figs 1 and 2. Only slight differences from the wild-type polyoma sequence were found. Py EC F9-1 has three insertions: a thymine between positions 85 and 86, another thymine after position 109 and an adenine between positions 119 and 120. One transition of a thymine into a cytosine at position 63 is also observed (The numbering system used here is based on the PyA2 sequence of Arrand *et al.*¹⁰ where residues in the late region are counted from the middle of the *HpaII*-3/5 taken as zero). In Py EC F9-5, there is a duplication of a 179-base pair region covering the origin of DNA replication but only the right-hand duplicated sequence reveals the existence of two supplementary nucleotides, a thymine inserted between positions 108 and 109 and an adenine between nucleotides 120 and 121. We do not believe that the duplication in Py EC F9-5 has any role in the properties of the mutant because other polyoma mutants with a duplication of the DNA origin of replication¹¹ do not give rise to productive infection on either EC F9 or EC PCC4 cells. The rest of the nucleotide sequence is identical to the one we have determined previously for the A2 wild-type polyoma¹², and differs only slightly from that published by Arrand *et al.*¹⁰. To confirm that the host-range capacities were due solely to the point mutations detected, and that no other modification elsewhere in the polyoma genome

Fig. 1 Genomic organization around the origin of DNA replication in wild-type and mutant polyoma, the DNA sequencing strategy. The origin of DNA replication is commonly given as the *HpaII*-3/5 junction. The beginning of translation of the early T antigens at -173 base pairs and the late VP2 at +291 base pairs, as well as its direction are shown by the corresponding arrows. Underneath are given the parallel regions of the two mutants analysed in the present study. The nucleotide sequence determined is presented by the underlying arrows. For both mutants, the *HpaII*-3, *HpaII*-5 and *BglII*-3 fragments were purified on polyacrylamide gels after digestion of the corresponding pBR322-Py plasmids, and labelled at their 5' end with ^{32} P using the exchange reaction with T_4 polynucleotide kinase as described by Berkner and Folk²⁰.

After secondary restriction with *Bam*HI, *Alu*I and *Bam*HI, respectively, the labelled and purified fragments were submitted to the partial chemical degradation of Maxam and Gilbert²¹ and the subsequent products were fractionated on 6 or 20% polyacrylamide-urea 0.35-mm thin gels²². Autoradiography was done either with Kodak NS-2T film at -20 °C or with an intensifying screen (Cronex Lightning-Plus, DuPont) using a preflashed Kodak X-OMAT R film at -70 °C (ref. 23). Point mutations observed are shown (▼). The duplication in polyoma EC F9-5 is designated by the box.



was involved, we constructed a hybrid DNA molecule in which the mutant *BglII*-*Bam*HI 750-base pair (930 base pairs in Py EC F9-5) fragment covering the mutated region was ligated to the complementing 4.5-kilobase region of wild-type polyoma, and vice versa. The hybrid DNA was used to transfect secondary mouse embryo cells and the virus burst was utilized to

infect EC F9 cells. Only the hybrids containing the 750-base pair (or 930-base pair) fragment from the mutants gave productive infection as assayed by the appearance of the T and V antigens (about 60–80% positive cells compared with 0.1% in the opposite construction). Segal *et al.*¹³ have detected a low amount of unspliced SV40 DNA in F9 cells after infection with

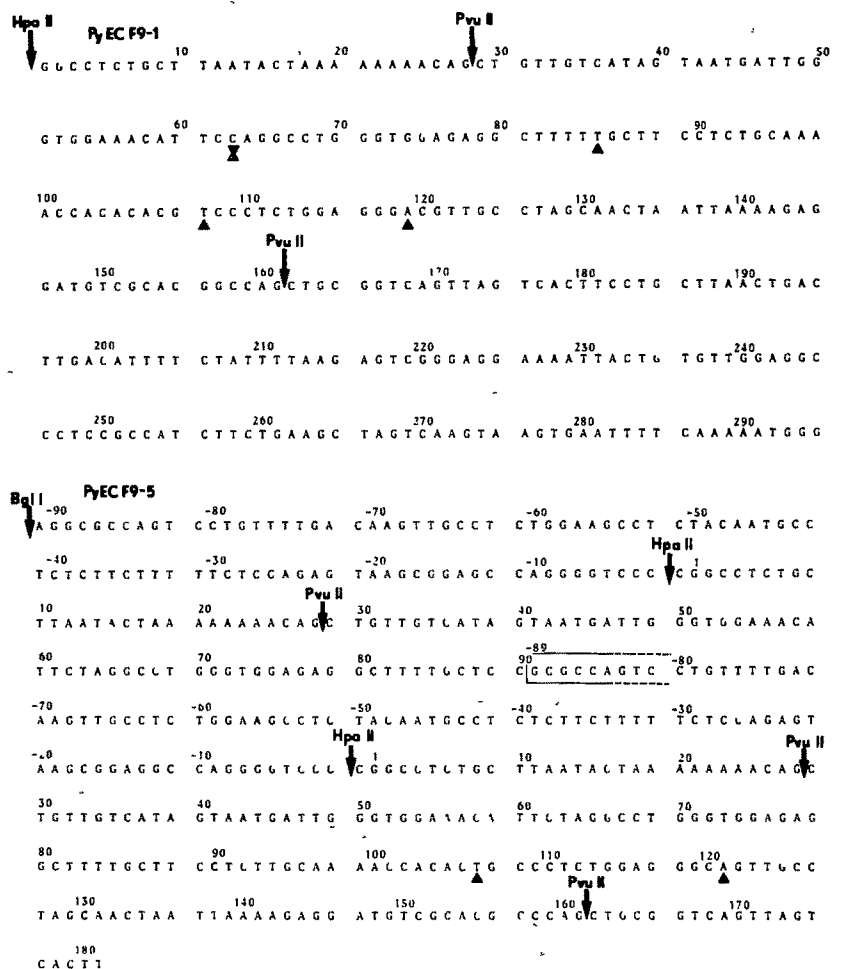


Fig. 2 Nucleotide sequence of Py EC F9-1 and of Py EC F9-5. The insertions (▲) and transitions (▼) in the mutant sequences are shown. The restriction sites for *BglII*, *HpaII* and *PvuII* are given as landmarks. The numeration of the nucleotides follows that of Arrand *et al.*¹⁰

this virus. The fact that the polyoma mutants are not located in the region concerned with the splicing of polyoma early RNA tends to exclude a possible block at this step during infection with wild-type polyoma.

The changes occurring in the Py EC PCC4 and Py EC F9 mutants show that both groups have genomes altered in the sequence of the *HpaII*-3 fragment, near the origin of DNA replication on the late side of the genome, in a noncoding region (Fig. 4). Py EC PCC4 mutants reveal a genomic rearrangement in which a deletion (from nucleotide 46 or 77 to nucleotide 107) is replaced by the duplication of a downstream region (nucleotide 138 or 157 to 220, see Fig. 4 and ref. 12), whereas the Py EC F9 mutants exhibit in the same region a series of point mutations as described above. Note that in the Py EC F9 mutants, one of the thymine insertions is always located at the 108/220 deletion-duplication junction site of Py EC PCC4 mutants. The region of the polyoma genome comprising nucleotides 62–133 can be folded into a clover-leaf structure (ref. 5 and Fig. 3). The generation of such a structure with circular double-helical DNA is certainly subject to topological constraints. However, transcription activation postulated for prokaryotic replicons¹⁴ or the action of a DNA gyrase (DNA topoisomerase II)^{15,16}, could permit the local melting of the DNA double strand and the formation of a secondary structure with single-stranded DNA. The structure can be stabilized by interaction with specific proteins. Alternatively, a secondary structure on the DNA level could be limited only to stem II (Fig. 3). It is striking that this part of the genome is modified in both Py EC PCC4 and Py EC F9. In Py EC F9, all the mutations are localized between nucleotides 63 and 121 (see Figs 3, 4) and touch mainly stem II of the clover-leaf structure, with the insertion of an adenine facing a complementary thymine. Such an insertion of an extra A-T base pair in a G+C-rich stem could change its interaction with proteins.

Whereas transcription *in vitro* with eukaryotic RNA polymerase II requires mainly the TATA box (D. Hogness and M. Goldberg, unpublished observations), several pieces of experimental evidence suggest that at least 100–200 nucleotides upstream of this sequence are important for transcription regulation *in vivo*^{17,18}. The rearranged region in the polyoma EC mutants is located 100–200 nucleotides before the hetero-

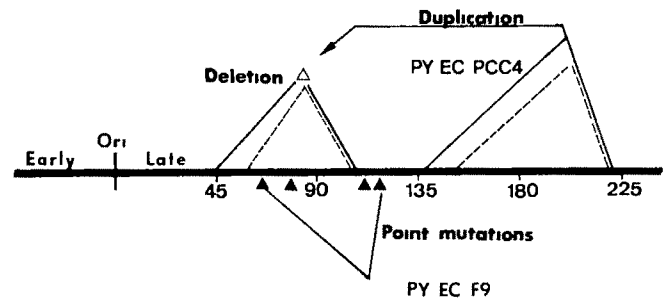


Fig. 4 Summary of the genomic changes around the origin of DNA replication in Py EC PCC4 and Py EC F9 mutants. The upper part of the diagram shows the rearrangements in the Py EC PCC4 mutants where a region between nucleotides 46 or 77 to 107 is deleted and replaced by the region between nucleotides 138 or 157 to 220 which is duplicated (see also ref. 12). The lower part concerns the Py EC F9 mutants with the point mutations (▲) between nucleotides 63 and 121 as described in the text.

geneous 5' ends of the early mRNA (R. Kamen, personal communication). In addition, this same region precedes the 5' ends of late mRNA and should be important for the control of late transcription. Thus, the sequence rearrangements that can occur without impairing the viability of the virus are limited. Previous studies have shown that wild-type polyoma DNA penetrates the EC cells but does not express early mRNA. Recent results¹⁹ reveal that on polyoma EC viral infection of PCC4 or F9, no new sites for early polyoma transcription initiation could be detected but the overall rate of specific polyoma transcription is greater than that of wild-type polyoma in these cells.

A model could be proposed in which a regulatory protein present in EC cells but not in differentiated derivatives inhibits early transcription of the virus by binding to the nucleotide 63–132 region. The modification of the sequence either by point mutations or by rearrangement destroys the specific binding site but still conserves a stable secondary structure otherwise needed for viral expression, thus relieving the inhibition of transcription. Alternatively, another model could be constructed based on a positive regulation. Interestingly, the putative tRNA-like structure in which the mutations occur is included in the region devoid of nucleosomes (gap) in viral chromatin (P. Herbomel *et al.*, unpublished results).

We thank C. Kress for assistance, C. Maczuka for help in the preparation of the manuscript, and A. Minty for valuable criticism of this manuscript. This work has been supported by grants from the French Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale and the Délégation à la Recherche Scientifique et Technique.

Received 11 November 1980; accepted 6 February 1981

- Swartzendruber, D. E. & Lehman, J. M. *J. cell Physiol.* **85**, 179–187 (1975).
- Boccardo, M. & Kelly, F. *Annls Microbiol. Inst Pasteur, Paris* **A129**, 227–238 (1978).
- Jakob, H., Boon, T., Gaillard, J., Nicolas, J. F. & Jacob, F. *Annls Microbiol. Inst Pasteur, Paris* **B124**, 269–282 (1973).
- Shermann, M. L. in *Teratomas and Differentiation* (eds Sherman, M. C. & Solter, D.) 189–205 (Academic, New York, 1975).
- Soeda, E., Arrand, J. R., Smolar, N., Walsh, J. E. & Griffin, B. E. *Nature* **283**, 445–453 (1980).
- Seif, R. & Cuzin, F. *J. Virol.* **24**, 721–728 (1977).
- Strickland, S. & Mahdavi, V. *Cell* **15**, 393–403 (1978).
- Pope, J. M. & Rowe, W. P. *J. exp. Med.* **120**, 121–128 (1964).
- Vasseur, M., Kress, C., Montreau, N. & Blangy, D. *Proc. natn. Acad. Sci. USA* **77**, 1068–1072 (1980).
- Arrand, J. R. *et al.* *J. Virol.* **33**, 606–618 (1980).
- Magnusson, G. & Nilson, M. G. *J. Virol.* **22**, 646–653 (1977).
- Katinka, M., Yaniv, M., Vasseur, M. & Blangy, D. *Cell* **20**, 393–399 (1980).
- Segal, S., Levine, A. J. & Khoury, G. *Nature* **280**, 335–337 (1979).
- Inokuchi, H., Dove, W. F. & Freifelder, D. *J. molec. Biol.* **74**, 721–727 (1973).
- Liu, L. F., Liu, C. C. & Alberts, B. M. *Cell* **19**, 697–707 (1980).
- Hsieh, T. & Brutlag, D. *Cell* **21**, 115–125 (1980).
- Grosschedl, R. & Birnstiel, M. L. *Proc. natn. Acad. Sci. USA* **77**, 1432–1436 (1980).
- Benoist, C. & Chambon, P. *Proc. natn. Acad. Sci. USA* **77**, 3865–3869 (1980).
- Dandolo, L. *et al.* (in preparation).
- Berkner, K. L. & Folk, W. R. *J. biol. Chem.* **252**, 3176–3184 (1977).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560.
- Sanger, F. & Coulson, A. *FEBS Lett.* **87**, 107–110 (1978).
- Laskey, R. A. & Mills, A. D. *FEBS Lett.* **82**, 314–318 (1977).

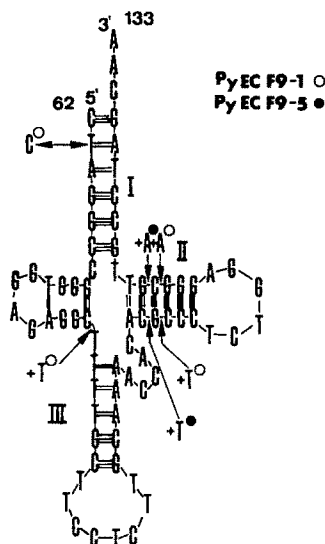
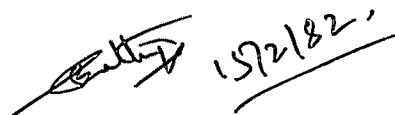


Fig. 3 A clover-leaf structure in the late noncoding region in the polyoma virus genome. The structure suggested by Soeda *et al.*⁵ for wild-type polyoma virus between nucleotides 62 and 133 on the late strand of the genome is drawn. It is possible to draw the complementary structure with the early strand. The insertions and the transition in this region in the two mutants analysed are shown. Note that most insertions for both mutants occur on stem II, always with a double insertion of the complementary adenine and thymine. Note also that in all the Py EC PCC4 mutants, the junction between the deletion and the duplication at 108/220 occurs again on the stem II¹².

30 April 1981



New guard at the National Academy

The annual meeting of the National Academy of Sciences in Washington this week is something of a milestone. Among other things, the scientific programme includes a meeting to mark the impending retirement of Dr Philip Handler as president of the academy. From then until 1 July, when Dr Frank Press takes over, he will be something of a lame duck, which does not mean that he will be less outspoken than in the past (President Eisenhower, one of the most taciturn and even inarticulate of presidents, waited until his lame-duck period before making his unexpected but still memorable attack on the "military-industrial complex"). Inevitably, however, the academy's army of committee members will from now on be guessing at the changes that Press's succession is likely to bring. Handler will have to try harder if he seeks to capture their attention.

He should not feel slighted. Handler has done a distinguished job in his two six-year stints as president of the academy. Neither of his immediate predecessors could have done as well. Dr Detlev W. Bronk's apparently randomly directed ebullience may have suited the spirit of the late fifties and early sixties, but would have landed him perpetually in hot water in the past decade. Similarly, Dr Frederick W. Seitz's masterly passivity, well suited to the booming budgets of the mid-sixties (but wearing thin towards the end) would have led him into trouble in the turbulent seventies. Handler's spell at the academy has been a much rougher ride. During this period, budgets have been falling, Congress has assumed much more explicitly than before the role of public watchdog over — some would say public prosecutor of — the scientific enterprise, which has also been threatened by the environmental nonsenses of the early 1970s and the fuss about genetic manipulation at the end of the decade. Inevitably, Handler has frequently been forced to play defending counsel. To his credit, he has done this job courageously — and in the process has made some stirring speeches.

Unlike the scientific academies of Western Europe, the United States academy has the seeds of dissension built into its constitution. Thus it resembles the academies of Eastern Europe in having a formal relationship with Congress, which is empowered to commission studies on whatever questions happen to flit across the scene. If, for example, Congress wishes to know whether small doses of radiation are dangerous, whether the accumulation of chlorofluorohydrocarbons in the atmosphere will dangerously deplete the ozone layer or whether the Department of Energy is telling the truth about energy forecasts, many committees of Congress will think first of commissioning a study from the academy. Technically, the work is done by the National Research Council, in whose management both the academies of engineering and medicine have a part, but in the public mind (and that of Congress) the academy of sciences is responsible. Since by definition the subjects on which Congress asks for advice are contentious, the academy frequently finds itself at the centre of public argument. Last year's report on the effects of small doses of radiation (see *Nature* 286, 550, 1980) was a good illustration of the troubles that can arise. Occasionally, the academy asks for trouble, as with its own committee's pungent advice that people should not be unduly concerned with the exact composition of what they eat (see *Nature* 285, 428, 1980). One of Handler's tasks, in the past twelve years, has been to give this stream of documents a collective stamp of authority. For the future, it would be better if the academy could arrange to shuffle off commissions for studies which are either pointless or untimely. Too much of what it has attempted should have been

refused, and should be in the future.

That is one task for Dr Frank Press. Another is to redefine the ground on which the scientific enterprise is defended. Hitherto, the academy's machinery for cogitating about problems of general concern has been directed towards specific contemporary problems, where data gathered by research promise to illuminate matters of general interest. The academy has been much less concerned with the machinery by which the scientific enterprise is sustained. The education of students in schools and colleges? The National Science Foundation has been allowed to set the pace, at least until its budget was slashed last month. The problem of university scientists, with or without tenure, engaged in the increasingly hazardous gamble of the pursuit of scholarship? Leave that to the American Council on Higher Education, which has strong opinions on issues such as tenure — and hope that people will one day recover from the over-abundance of documents churned out by the Kerr commission on higher education. The propriety of the arrangements academics make with their universities to share in the profits of extramural commercial enterprises? Or that universities make with commercial companies or with agencies of the federal government? The universities, the argument seems to go, are too jealous of their independence to welcome advice.

As events have turned out, the academy in Washington should move quickly to take up these and a host of related questions. Although the competitive universities and colleges of the United States are rightly proud of their autonomy, they are now afflicted by so many common and pressing problems that the academy's influence could be decisive. Should the protest against the new strict interpretation of the rules prohibiting the export of military material, which threatens to restrict the education of overseas graduate students, be left to a handful of college presidents (see *Nature* 9 April)? Are the new rules introduced by the Office of Manpower and Budget for accounting for the use of federal funds equitable or merely iniquitous? And will they achieve their declared purpose of saving federal dollars? Is it sensible that sources of financial support for graduate students should be as much curtailed as in the past few years? What — in the national interest as well as that of research — should be the balance in the funding of research between project grants (with their accompanying variable overhead percentages) and straight institutional support (of which there is at present almost none)? And if institutional support comes back into fashion, to which institutions should it be channelled by which agencies? Many of these are tedious questions, but they are much more important than the consequences of aerosols in the atmosphere.

The surprise so far is that the academy seems to have had very little interest in these issues. They are of course intangible and often boring. There is also always the danger that if a body such as the National Academy of Sciences speaks out forthrightly on these subjects, it will be accused of self-serving propaganda on behalf of its members. That risk is worth running. Indeed, the risk must be run. For with so many other well-wishers of the scientific enterprise emasculated by President Reagan's first budget, the academy is almost the only institution in Washington able to speak out. Moreover, the issues are likely to induce a harmony of purpose among the members of the academy sufficiently strong to carry the institution through the divisive issues certain to crop up in the next six years as in the past twelve — whether, for example, to protest at Sakharov's exile and, if so, how? Is this an agenda for Dr Frank Press?

Will the London Zoo survive (and how)?

The London Zoo, more formally known as the Zoological Society of London, is in serious trouble. In the annual report for 1980, published last week, the president of the society, Lord Zuckerman, says that the organization is "up against a brick wall". The evidence of the report leaves little doubt that he is correct. Last year, the society made a loss of £666,000, more than 13 per cent of its income of just under £5 million. This, however, is not the end of the tale. The society has made large losses in previous years — £130,000 in 1979, for example. The novel but ominous implication of last year's result is that the society's bank balance has been almost emptied, and replaced by a substantial overdraft. And that there are unmistakable signs that the chief source of revenue, the entrance fees paid by visitors to the two zoos, in Regent's Park, London, and at Whipsnade, is no longer an inexhaustible pot of gold. For at least two decades, as charges have been increased, the numbers of visitors to the zoos have been falling. Last year, attendance at Regent's Park fell by 11 per cent. Like many other enterprises in inflation-ridden Britain, the society (which is a charity and not technically a business) is being crippled by continually rising costs and the shortage of cash in the economy to spend on the services it provides.

The society's recipe for survival, outlined in the report for 1980, is logical enough but almost certainly destined for failure. The zoo, the argument goes, is constituted as a charity which provides public services of several different kinds — pleasure for visitors to the two zoos, educational experience for a great many school students and a substantial amount of scientific research, much of it distinguished and not all of it inspired by the veterinary needs of the animals on display. Last year, the two zoos managed to collect some £60,000 on account of their overtly educational work, by almost any test a miserly recompense for their educational value not merely in London but in the United Kingdom as a whole. Last year, the cost of research (£780,000) exceeded by £322,000 research grant income. So the society is asking why its value as an educational institution should not be more tangibly recognized by local and central government, and why a substantial part of the cost of its research must continue to be met by those who innocently spend an afternoon looking at exotic animals. On both counts, the zoo has a powerful case.

If the zoo were to go out of business — and if its financial position does not improve, it may have no choice — there would be a great outcry from all kinds of interested parties. Local education authorities would be prominent among those protesting that a great institution should not be allowed to sink without trace. They would be complaining that children in their schools would be deprived of an essential part of their education. The spendthrift among them would no doubt argue that if the London Zoo should be closed, it would be necessary for public funds to pay the cost of carting children to the distant game parks of central Africa. Some might even be able to raise the funds. The zoo's case is that if the local authorities are as certain as seems likely to mourn the disappearance of the London Zoo (and perhaps also of Whipsnade), they should contribute more generously now to help keep both institutions in being. The logic is impeccable. The snag is that local authorities are just now as hard-pressed as the zoo itself. And in the nature of British local government, they are incapable of making discretionary expenditure, however prudent. Voluntarily to shoulder part of the cost of supporting the London zoos would be politically risky, but the same local governments (the Greater London Council and the Inner London Education Authority) would have no qualms in foisting on their ratepayers the much larger cost of salvaging the society's collapse.

The obstacles to the zoo's case for more outside support for its research programme are a little less ominous. The point has already been established that for two years there should be a modest subvention (£134,000 last year) from central government funds, while the zoo's laboratories can compete for project grants

from public and private sources. The accounts for 1980 (which sensibly charge the cost of research "of direct benefit to animal management" against the income of the public exhibits) nevertheless show that the society had to spend £250,000 of its inadequate income on research. The principle that a private organization should qualify for a public subvention towards the overhead cost of research is unusual but is not an outrage. If a research programme is worthwhile, there is no reason why it should not qualify for public support as if it were carried out in some publicly owned laboratory. The obvious snag, that private management may mean fickle management, is easily overcome by making sure that all grants are short-term grants. In the zoo's present plight, the Advisory Board on the Research Councils, which unimaginatively cuts the annual research cake in Britain, should be willing to do more to help.

Public support of this kind, however, can only be a stop-gap. The zoo itself must acknowledge that its research institutes cannot prosper if they have to live from hand to mouth, keeping watch on the dwindling queues at the visitors' admission gates. It is now clear that the zoo's laboratories have a modest if distinctive contribution to make to some intriguing problems in animal physiology and veterinary medicine, but that they will never be comprehensive or dominating. In the interests of the laboratories, the zoo should therefore seek a more durable framework for its research. For a year now, there has been talk of a consortium including the zoo, the British Museum (Natural History) — which should simplify its name as well as root out cladistic excesses from its public exhibitions — and the Royal Botanic Gardens at Kew. All three institutions have in common their dual concern for public exhibition and scientific research. The connections between the three research programmes are by themselves, however, probably too tenuous to be a durable bridge. Nor is it likely, in present circumstances, that other British zoos would be willing to regard the London Zoo as a central research service for them all, writing cheques accordingly. The zoo's research enterprise might be better able to make its way in the world by setting itself up as a common service for academic and public research laboratories, for example by setting up well-run breeding colonies of primates. Either (or some other) way, the zoo must quickly work out a plan and set out to make it work. In the process, it may find it prudent to put more distance administratively between the research laboratories and the Zoological Society of London as a whole. There is no evidence that the fellows of the society who go to watch the pandas on Sunday mornings presume to interfere with the research, but these days a show of independence would do no harm.

Much the same would be necessary if the zoo were to succeed in recruiting support from local government to keep its zoological display in good shape. Local government politicians are more convinced than others that money means power, in this case influence over exhibition policy. Lord Zuckerman seems bent on persuading London's local politicians that they should follow the examples of their peers in Berlin, Rotterdam and elsewhere. Would he and the council of the zoo be prepared to share their present absolute power with those willing to keep the zoo out of the red? This unfamiliar thought is neither unthinkable nor unseemly. In the long run, it would help and not hinder if local governments were prepared to share anxieties about the zoo as well as support for it. Experience elsewhere suggests that the propriety of the display need not be damaged. The implication is that the zoo should evolve into a consortium of three separate charities — a research organization, a scientific society of some importance and a public exhibition. In these hard times, it is too dangerous to let the enterprise stand or fall as a whole. That way, it would be better organized for survival, better able to provide its well-wishers with a sense of belonging to a constituency. After all, many of the society's friends think it wrong that animals should be kept in zoos.

Vienna institute hit by espionage charges

Claims force Soviet official to resign

Allegations of espionage for the Soviet Union have forced the resignation of Dr Arkady Belozеров as secretary of the Institute of Applied Systems Analysis at Laxenburg, near Vienna. This dramatic development has intensified anxiety at the institute about its long-term financial security, already threatened by the Reagan Administration.

Dr Belozеров was appointed to his post at the institute in 1971, and took up his job in December that year. An article in the British news magazine *Now* on 10 April said that Dr Belozеров, a plasma physicist by training, was also an agent of the KGB who had been the Soviet contact with a Norwegian double agent during his spell in Vienna. A statement put out by the institute on 15 April said that Dr Belozеров had offered his resignation and that this had been accepted.

The responsibilities of the secretary of the institute are largely those of external relations with member organizations (mostly national academies), potential donors (such as foundations) and the press. As with other senior posts at the institute, the secretary is appointed jointly by the director and the governing council. Dr Belozеров's name was the only nomination before the council in 1979. The institute's policy on recruitment requires that there should be a national balance among senior members of the staff.

The allegation in the *Now* article is that Belozеров took over from another Vienna-based Soviet agent responsibility for keeping contact with a Norwegian petroleum engineer known as Jansen. According to the magazine, "Jansen" began working as a double agent for the Norwegian security services six years ago, providing a succession of Soviet contacts in Oslo and then Vienna with innocuous and sometimes misleading information about petroleum operations in the North Sea and the China Sea.

Dr Belozеров is reported to have met "Jansen" on several occasions in Vienna, to have been incompetent in making arrangements for meetings and to have been reluctant to pay for information provided until its authenticity had been checked. He is reported to have given a full account of his career as a member of the KGB and in the process to have revealed the identity of a Soviet agent in Austria. His meeting with "Jansen" arranged for 4 April outside the Volks Opera in Vienna is said to have been frustrated by the expulsion of another alleged Soviet agent

from Norway.

In his letter of resignation to the institute, Dr Belozеров protests that the allegations "are without factual basis", but that there is no impartial forum in which he can demonstrate them to be false. His letter says that he offered his resignation "in the interests of the institute".

The institute says that Dr Belozеров's conduct of his post as secretary had been "completely satisfactory" but that it has no way of independently determining the truth of the allegations against him and therefore had no choice but to accept his resignation. A spokesman said last week that in the present precarious financial circumstances, any member of the staff against whom allegations of this kind were made would be asked to leave.

One curious feature of the Belozеров affair, as reported, is that Soviet agents should have been willing to pay for information about oil technology in deep waters

such as those of the North Sea when the institute itself had only recently published an up to date account of the field.

The Austrian government, an enthusiastic benefactor of the institute, has reacted protectively. The Minister of the Interior, Mr I. Lanc, said last week that Belozеров had committed no offence under Austrian law and complained that "certain countries" not having common frontiers with the Soviet Union had come to undervalue détente and to regard the Institute of Applied Systems Analysis as an expendable target.

The institute's financial prospects are not nearly as bright as they may have seemed only a few months ago. Although the member organizations of the institute are national academies, most of these are reimbursed by their governments for the cost of membership. The Soviet Academy of Sciences and the United States National Academy of Sciences are the largest con-

Schuster quits as EEC research director

Brussels

The reshuffle of top jobs at the European Commission following the appointment of commissioners and the allocation of their portfolios at the beginning of 1981 has led to a flurry of resignations during the past two weeks, and rumours that some directorates-general are to be reorganized are gathering force.

Five directors-general are known to be leaving, among them Gunter Schuster, head of Research, Science and Education (DG XII). His resignation was handed in amid speculation that he is dissatisfied with the way responsibility for DG XII has been split among three commissioners. At the beginning of the year, infighting among the new 14-man Commission caused a messy distribution of the research portfolios. Etienne Davignon (France) grabbed research and science, leaving Ivor Richards (United Kingdom) with education and training, and Karl-Heinz Narjes (West Germany) with industrial innovation.

Some members of the Commission claim that this arrangement is not as awkward as it might seem — decisions taken by the commissioners have in any case to be unanimous. The real reason for Schuster's departure may be that at 62 he would be due for retirement mid-way through the next four-year indirect action programme now being planned. A new man would be able to see the plan to completion.

His departure leaves a vacuum hard to fill. According to a close colleague, Herbert Allgeier, Schuster's greatest personal achievement was to build up from scratch the European Community's research policies from the shambles that followed the collapse of the Euratom programme at the end of the 1960s. Pressure is now being brought to bear on Davignon to

find a replacement as soon as possible, so that the new appointee can help draw up the new research policy guidelines.

Among the names being mooted is that of Fabrizio Caccia Dominioni, at present looking after research, development and nuclear policy. However, somebody may well be brought in from outside, although probably not another German or Briton. Leonard Williams of the



Schuster leaves the scene

energy directorate-general is being replaced by Christopher Audlands. In the best Byzantine traditions, the top jobs are distributed so that national pride and influence are not diminished with each four year reshuffle.

Other changes are rumoured, but one decision, it seems, has definitely been taken. The Environment and Consumer Protection Service (ECPS) is to be elevated to directorate-general status thanks partly to the patronage of the European Parliament. The ghostly DG XI, which was left vacant after an earlier round of administrative musical chairs, will once more have an occupant. Michel Carpentier will remain in charge as head of ECPS and will henceforth be officially addressed as director-general.

Jasper Becker

tributors at rather more than £1 million each a year

In the United States, the National Academy's annual contribution has since the inception of the institute been met by the National Science Foundation from its allocation of funds for international activities, cut in President Reagan's first budget from \$15 million to \$10 million for the coming financial year. After protests from the academy that the statutes of the institute require twelve months' notice of withdrawal, the putative contribution for 1982 has now been restored to the foundation's budget, but on the strict understanding that this will be the last contribution to be sanctioned by the Administration.

In principle, there is no reason why the academy should not seek to raise its annual contribution from some other source than the United States government, but in present circumstances this will not be easy. The 1982 budget for the institute will be agreed between the member institutions at two council meetings to be held in June and November this year. The United States academy will have to give notice of intended withdrawal at the end of this year if it is not by then assured of funds with which to pay its 1983 subscription.

It is ironical that the institute, founded partly with the intention of providing a place at which people from East and West could work on problems of international importance, should have run into trouble because of allegations of espionage.

US graduate employment

Industry has jobs

Washington

Despite continued uncertainty in the US economy, business is booming for engineering graduates — and natural scientists are not far behind. Figures just released by the College Placement Council of Bethlehem, Pennsylvania, reveal that engineers who expect to graduate this summer are already being offered starting salaries of up to \$26,000, a 10 per cent increase over last July. Graduates in physical and Earth sciences were being offered \$22,000, an 18 per cent rise compared with last year.

This rapacious demand for technical graduates has been confirmed by a report from a New York consulting firm, Deutsch, Shea and Evans, which claims that there has been a sharp up-swing in employment prospects for graduate engineers since last November. And the figures are in marked contrast with starting salaries for arts graduates, which average about \$12,000, with many failing to find any jobs at all.

The shortage of engineers is partly a reflection of declining enrolments in university engineering schools. For example, 2,751 engineering doctoral degrees were awarded in 1980, compared

with 3,074 in 1972, although this also reflects the fact that fewer engineers are staying on to do graduate work, the 1980 figures include more than 60 per cent foreign students, many of whom were not expected to stay in the United States.

Recruitment difficulties have also produced near-crisis in the engineering schools, few of which can afford to match salaries in industry. According to William A. Cox, president of the National Society of Professional Engineers, there are now 1,800 unfilled faculty positions in engineering schools — a situation which has led the society to add its voice to those asking Congress for federal assistance to upgrade facilities and research equipment to make the educational environment more attractive.

There is also a feeling that the turnaround in demand for technical graduates reflects greater optimism in the economy than six months ago. "In general, the market in technical areas is the best that it has been for years", says Victor Lindquist, director of the placement office at Northwestern University which carries out regular surveys of employment prospects. He claims, for example, that metal companies such as mining and steel, sectors which are likely to benefit directly from a significant increase in military spending, seem to be stockpiling talent in anticipation of a quick recovery in the economy.

The College Placement Council survey revealed that, as in the past, engineers represent the hottest buy in the employment market. More than 63 per cent of the job offers reported in the survey as having been received by mid-March were to future engineering graduates, although these only made up 7 per cent of the bachelor degree graduates in 1980-81.

In proportional numbers of offers, the scientific disciplines were dominated by computer science graduates, another field where demand is considerably in excess of supply, a point which has been made recently by the National Science Foundation. However, the average starting salary offered to such graduates of \$19,968 was considerably lower than the \$21,912 offered to graduates in physical sciences.

Among masters level graduates, the highest offers went to chemical engineering graduates, with an average starting salary of \$26,240 — almost 13 per cent higher than the figures last July. Geology and related geological sciences, in strong demand in the petrochemical and mining industries, ranked second with an average salary of \$25,380. All these figures are based on offers reported by placement offices at 161 US colleges and universities.

Figures released last week by the National Science Foundation reveal that there is an increasing tendency for graduate scientists, engineers and technicians to enter service rather than manufacturing industries. These industries accounted for 60 per cent of the total growth of employment for graduates in the three categories

between 1978 and 1980, even though only accounting for 40 per cent of the total employment.

Total science and engineering employment in private industry was virtually unchanged between 1970 and 1980, but this trend may be reversed, however, if the current patterns of employment for this year's graduates is maintained.

David Dickson

Soviet space programme

No shuttle yet?

The high cost of space research is troubling Soviet planners as well as their counterparts in the space agencies of Europe and the United States. At a ceremony on 10 April commemorating the twentieth anniversary of Yuri Gagarin's first manned space flight, the president of the Soviet Academy of Sciences, Dr Anatoli P. Aleksandrov, spoke in a low key for such an occasion.

Aleksandrov's speech was little more than a justification of expenditure so far, and a defence of the proposition that the more that is spent on space, the greater the eventual benefit. This is not the first indication that Soviet space research may be running into financial difficulties. Earlier this year, the Soviet side had to make it clear to their French partners that if the two countries were to carry out a joint mission to Halley's comet, spending on their Venus mission of 1986 would have to be halved.

To demonstrate the priority being given to cost-effectiveness, Aleksandrov explained that three and a half years ago the cosmonauts aboard Salyut-4 were working on behalf of only 40 scientific and industrial bodies, whereas in the latest mission, Vladimir Kovalenko and Viktor Savinykh on Salyut-6 were providing data for some 400 organizations.

Naturally Aleksandrov did not discuss any military aspects of the Soviet space programme, although he did add to the abuse of the US space shuttle as a military vehicle. The press in the Soviet Union suggest that the US shuttle is primarily for launching laser weapons into space.

Soviet intentions on reusable launching systems remain unclear, however. Aleksandrov did not discuss any such plans, but Soviet cosmonaut Vitalii Sevastyanov stated in the Slovak daily *Pravda* that the Soviets were considering the construction of a shuttle. He claimed, though, that the present conventional methods used in Soviet space launches were more convenient than a shuttle for their programme based on large orbital stations designed for "peaceful purposes" and "suitable for a longer stay in orbit of scores of specialists".

There is speculation in the West, however, that at least one experimental Soviet shuttle has been launched — number 929 in the "Kosmos" unmanned satellite series.

Vera Rich

Anglo-German project Aurora held up

Red tape has already delayed the British half of an Anglo-German programme of auroral investigations for several months. Last week, one of the obstacles — agreement on the wording of a lease for land at a site in the north of Scotland — was surmounted. But a further six months delay is now likely, pending the installation of an electricity supply to the site.

The joint programme, known as SABRE, is for radar backscatter observations of the aurora from two sites — near Wick in Scotland and at Uppsala in Sweden. The two collaborating groups are led by Professor Tudor Jones at the University of Leicester and Professor I Axford at the Max Planck Institut für Aeronomie, Lindau. The British Science Research Council and the Max Planck Institut are supporting the experiments to the tune of about £200,000 each.

The programme is designed to complement a similar twin auroral radar, called STARE, operated from two sites in northern Scandinavia by the Lindau group. Observations from the Swedish site began last January, but delays in preparing the Scottish site have meant that spatially and temporally resolved observations have not yet been possible.

The Scottish site is one of four sites originally owned by the Ministry of Defence but now controlled by the Property Services Agency which deals with matters relating to public land. Most of the hold-up seems attributable to delays in reaching agreement on improving accommodation and connecting an electricity supply. The agency, which plans to sell the other three sites, has been concerned that the electricity supplies of each should be independent. At present, the supply to the site which the University of Leicester wants to lease is routed via a transformer situated elsewhere. Four options for installing a new supply, presented by the North of Scotland Hydroelectric Board, have been turned down as too costly.

Some blame the delays on the snail's pace at which the Property Services Agency works. After 18 months of turning the bureaucratic handle, an acceptable draft lease finally arrived at the University of Leicester only last week. But the problem of installing a new electricity supply remains. A mains cable supplying the site was apparently damaged a few years ago when a new road was built. Repairs involve laying new cables. The North of Scotland Hydroelectric Board estimates that the job could take six months and cost about £5,000. The agency has now handed over responsibility to the university.

If work cannot be speeded up, a temporary generator will have to be installed if experiments are to begin before winter.

Judy Redfearn

US nuclear power Cost disputed

Washington

Efforts to meet the safety problems created by nuclear power have already made nuclear energy considerably more expensive than electricity from modern coal-fired plants, even with the sophisticated anti-pollution equipment needed in coal-fired stations. And the gap is likely to increase at least until the end of the decade according to a report published in New York last week by an independent energy economist, Charles Komanoff.

The nuclear industry has reacted sceptically to Komanoff's prediction, claiming that he underestimates the productivity of nuclear power, and overestimates the productivity of coal generation. The Atomic Industry Forum, chief spokesman for the nuclear industry, issued a statement saying that the conclusions clash with those of other economists, and that "to predict cost trends for coal and nuclear in a period of fuel-price volatility and still-uncontrolled inflation seems highly dubious".

However, Komanoff, who last year gave evidence on nuclear power costs to the UK House of Commons Select Committee on Energy, is standing by his figures. He points out that similar scepticism greeted a prediction he made five years ago that the nuclear industry would be operating at only 55 per cent of capacity through to the end of the 1970s, a figure which has turned out close to the truth.

He also claims that the argument about coal being cheaper than nuclear power appears to be supported by figures published by the Atomic Industry Forum itself. In a report issued two months ago, the industry group estimated that electricity from the three nuclear plants which started operation in 1978 cost 2.5 cents per kilowatt hour to produce, compared with 2.0 cents for coal and 5.7 cents for oil.

Komanoff maintains that the more nuclear power has expanded, the more generic problems it has revealed, and the greater the effort required to prevent the number of nuclear incidents from increasing at the same rate. He says this is substantiated by the steady increase in the number of warnings of potential generic problems by the Nuclear Regulatory Commission during the past decade.

Komanoff bases his future projections on the cost of obtaining electricity from coal-fired plants by assuming that increasingly stringent controls will require the use of scrubbers to remove sulphur dioxide — devices that can consume up to 25 per cent of the capital costs of a new plant — and that the cost of coal will increase at a rate about 2 per cent more than the inflation rate. On this basis a coal plant will increase in capital cost from \$583 to \$794 per kilowatt hour of generating capacity in "constant" dollars. In

contrast, extrapolating on past trends in the rising costs of nuclear power plants due to the need to include increasingly stringent safety precautions, Komanoff predicts that the cost will rise from \$887 to \$1,374 per kilowatt hour of capacity.

The Atomic Industry Forum is looking closely at Komanoff's figures. Its own analysis, based on power plants coming into operation between 1970 and 1978, is that "nuclear electricity in 1979 continued to show a modest advantage in total generating cost compared with neighbouring coal units", and that nuclear power also compared favourably with coal in capacity factor. At the same time, the industry group admits that regulatory delays and public intervention in licensing procedures "run up nuclear costs significantly", it claims that there are approximately one dozen nuclear plants currently snarled up in bureaucratic delays which will cost US consumers as much as \$3,000 million in higher utility bills.

Whichever analysis of the relative costs of nuclear and coal is correct, however, both sides agree that immediate prospects for the US nuclear industry are bleak. Only 13 new nuclear-power plants have been ordered in the United States since 1975, and plans for 50 plants have been cancelled — many because of decreasing predictions of future energy needs.

On top of this, no new plants have been ordered in the past two years. The nuclear industry is hoping that President Reagan will make good his election promises of increased support, and is already providing him with suggestions as to how this might be achieved, from revamping some of the safety regulations which Komanoff claims to be at the root of the cost escalation, to proposals for federal financial aid for nuclear-plant construction.

Overhauling the regulatory apparatus, however, will not be easy. Only one of the five seats on the Nuclear Regulatory Commission is vacant, and two of the existing commissioners have been harsh critics of some of the industry's demands. In addition, Mr Reagan's rumoured choice for the chairmanship of the commission, former head of naval operations Admiral James L. Holloway, is reported to have turned down the offer, and with no clear indication of the direction in which it is meant to be moving, the commission is running at half speed.

David Dickson

Anti-viral drugs

Herpes on trial

Volunteers, 50 male and 50 female, with genital herpes simplex infections are being sought by the Infectious Diseases Unit at the University of Texas Health Science Center at Dallas to take part in clinical trials of the promising antiviral drug acyclovir (acycloguanosine). There is as yet no effective treatment for genital herpes, which is sexually transmitted and now

considered to be of epidemic proportions in the United States, where 300,000 new cases are reported each year. Each new patient becomes a carrier for life. Acyclovir has generated great enthusiasm at Burroughs Wellcome, its developers, as a real advance in treating the disease. One worry, though, is that viruses may acquire resistance.

Herpes simplex virus (HSV) can infect any part of the skin or mucous membrane. HSV Type 1 cause the familiar "cold sore" blisters on the lip. Genital herpes, caused by HSV Type 2, is a more severe condition and causes painful blisters in the genital area. In the Dallas trial the drug will be applied as an ointment over skin areas where patients feel a tingling, burning or itching sensation (a "prodrome") suggesting that a blister is about to develop. Acyclovir will be compared with placebo cream containing no active agent. If the drug succeeds in reducing the period of recurrences to one or two days, it should not only reduce the period of pain for the patient, but also reduce the capacity for passing on the disease to others.

Because viruses are parasites of the normal systems of the cell they infect, antiviral drugs tend also to act against the host cell. There are few points at which the life processes of the virus can be properly differentiated from those of the host, but acyclovir acts specifically against herpesvirus through one such reaction (*Nature* 272, 583, 1978). It is phosphorylated by the herpes-specified enzyme thymidine kinase, so that the active triphosphate derivative is present in much greater quantities in infected than in uninfected cells. In a sense, the virus thereby invites its own destruction. It remains to be seen whether acyclovir will prove active against even the dormant form of the virus. Herpes infection is characterized by the way in which all outward evidence of the infection can disappear, but the virus persists undetected in certain nerve cells.

Acyclovir is being used as an ointment in the Dallas study, but later tests will include systemic (for example intravenous) injections which could increase the range of infections against which the drug can be used. Already acyclovir has proved active as an eye ointment against herpes simplex corneal ulcers, and less conclusively against viral pneumonia.

The US National Institutes of Health are to fund trials of intravenous acyclovir in newborn babies with neonatal herpes, in comparison with adenosine arabinoside (vidarabine), and in cases of herpes encephalitis, in comparison with the adenoside of adenosine monophosphate. Neonatal herpes is a herpes simplex infection contracted from the infected mother during passage through the birth canal — if the virus reaches the brain it can cause death. Herpes encephalitis is caused by reactivation of the virus in the brain.

Charles Wenz

European Earth sciences Unity prevails

Strasbourg

Europe's Earth scientists may yet unite. A meeting on 13–16 April at the assembly of the Council of Europe saw the first public appearance of the European Union of Geosciences (EUG). At present, EUG consists of a founding committee and has the official status of a working party of the Parliamentary Assembly of the Council of Europe. Its aim is to provide an organization which encourages the exchange of ideas and information across the more traditional subject barriers that separate geologists from the geochemists, geophysicists and oceanographers. To that end, membership is open to any geoscientist, but officials of EUG are to be nationals of the member countries of the Council of Europe.

At present, EUG is run by a founding committee under Professor C. J. Allegre of the University of Paris, and its formal statutes are to remain provisional until approved at a future meeting. EUG intends to hold a regular meeting in the spring in Strasbourg every other year, its interdisciplinary scope is to be complemented by more specific topical conferences to be held between the more general meetings. The concept and format of EUG owe a lot to the success of the American Geophysical Union and its big annual spring and fall meetings, which have almost become a must for many European researchers. However, the scope is to be wider insofar as it is to encompass all Earth scientists, not only those in the geophysical community. In this way, it hopes to avoid rivalries and misunderstandings with the European Geophysical Society, which caters for the same scientific community as the American Geophysical Union.

Although the European Science Foundation (ESF) did not play a major part in the organization of the Strasbourg meeting, it appears that ESF is prepared to help set up and support a small permanent office for EUG in Strasbourg. Is this the beginning of ESF taking a more active interest in the Earth sciences, an area hitherto neglected by its working parties?

The European Parliament building was made available by the Council of Europe for this first EUG meeting. Although organized at relatively short notice, it nevertheless managed to attract more than 600 participants and enough presentations to run a full four-day programme grouped in 16 different sessions.

The abstracts from the first meeting appeared as the first issue of an ambitious new publication called *Terra Cognita*. EUG hopes to develop this into a quarterly journal which should serve as a forum for scientific information on research projects and activities as well as to promote the discussion and exchange of ideas within the Earthsciences community. Konrad Guettler

Out on a limb

Brussels

Dutch Socialist Member of the European Parliament, Annie Krouwel-Vlam, is trying to whip up support for a European Community directive on common rules for organ transplants. She is surprised that there are no rules regulating relations between organ banks and facilitating access for would-be recipients of new parts. Such measures would implant a greater awareness of the benefits of the European Community in the breasts of its citizens, she asserted.

Her suggestion has been passed to the Community's organ banks and the European Commission. So international banking could soon take on a new meaning.

Jasper Becker

Interferon business

Trials ahead

Rotterdam

That interferon is big business can only have escaped the notice of those totally occupied for the past two years in preventing the pollution of space by small plastic tiles. That it continues to be big business was evident from the number of companies represented last week (19–24 April) at the interferon meeting organized by TNO, the Netherlands organization for applied scientific research.

Half of the British participants were from Beechams, Celltech, Fisons, ICI, G. D. Searle and Wellcome, with about the same proportion of commercial people among the West Germans. The large United States contingent included representatives of Cetus, Du Pont, Enzo Biochem, Electro Nucleonics, Genentech, Hoffman-La Roche, Hydron, Meloy, Sepetex, Schering and Shell. Not a slide went unphotographed nor a talk unrecorded by one or more commercially interested party.

Many of these companies already had a declared interest in interferon but some were new to the game. ICI let it be known that it has cloned an interferon gene from lymphoblastoid cells in collaboration with Dr E. de Maeyer's group at the Institut Curie in Orsay. ICI says that this is just an exercise in cloning, but with the setback in the fortunes of the company earlier in the year (*Nature* 12 March), there must be thoughts of trying to cash in on clones.

There is little doubt that the demand for highly purified interferon is growing, but the growth can be fitful. Wellcome, for example, recently claimed that supply had outstripped demand for its alpha-interferon (from white blood cells). That situation could easily be reversed if the current clinical trials of Wellcome's interferon show promise. The interim results, reported in Rotterdam, of several relatively large trials of a variety of different alpha-

interferons suggest that it has a tantalizing degree of activity against a broad spectrum of viral diseases and tumours. But activity is not enough, or consistent enough, in most diseases to give a final verdict without more trials. Also, as genetically-manipulated bacteria start to yield sufficient amounts of the ten or so subtypes of alpha-interferon, the permutations of subtypes and diseases becomes huge.

Meanwhile, large-scale trials of beta-interferon (from fibroblasts) lag slightly behind. The largest has yet to start. Financed and organized by the National Institutes of Health (NIH), the trial will use 50,000 million units of beta-interferon, enough for several hundred patients. The interferon is being supplied by Flow Laboratories Inc, which won the \$2 million contract a year ago. First deliveries were due a few weeks ago, but NIH appear to be in no hurry. This is lucky for Flow Laboratories, whose own production has not been going smoothly. Fortunately, they have a subcontract with the West Germany company of Dr Rentschler Arzneimittel GmbH which has continued to amass beta-interferon made by its own, less advanced, techniques.

The next big contract, also for NIH, will be for gamma-interferon (from cells of the immune system). Far less is known, at least publicly, about this type of interferon, but Biogen and Genentech are rumoured to be on the verge of cloning the relevant genes.

Other commercial interests centre on monoclonal antibodies against the interferons. Several new monoclonals against either alpha or beta-interferons were announced in Rotterdam, but only the original monoclonal (see *Nature* 285, 446, 1980) is as yet commercially available. The sellers are the British company Celltech Limited, so far its only product. It will be surprising if the Celltech catalogue does not double soon.

Peter Newmark

Dioxin hazards

Secrecy at Coalite

Secrecy about the medical consequences to workers exposed to dioxin continues at the British company Coalite and Chemical Products Limited. Documents now to hand show that Coalite will not agree to the further medical investigation of workpeople if the physicians concerned insist that the results are eventually published.

Coalite's shyness is not unprecedented. Last year, the company withheld information on an epidemiological survey of workers exposed to dioxin (2,3,7,8-tetrachlorodibenzodioxin) at its Bolsover (Derbyshire) plant in 1968-71 (see *Nature*, 6 March 1980). The exposure followed an explosion in a vessel used for the manufacture of 2,4,5-trichlorophenol and the workers affected by dioxin developed the skin disease chloracne.

The fresh trouble which has arisen at Coalite results from the need for more

detailed investigations of individual workers to make up for the deficiencies of the epidemiological survey of 1977-78. That involved only 41 of the 90 workers known to have been exposed, but revealed (almost a decade after the event) some abnormalities of blood chemistry and immune function — but none of the expected chromosomal abnormalities in circulating lymphocytes.

The news that Coalite was unwilling to disclose information about this survey served to alert the statutory Employment Medical Advisory Service of the Health and Safety Executive to the existence of such data, which was disclosed on a confidential basis after three weeks. The service has also persuaded Coalite that its medical adviser, Dr George May, should publish some of the results.

The employment service has in the meantime been uncomplimentary about the survey which, it says, is deficient in its failure to follow up past employees and in the inadequate matching of the control populations. The survey covered 41 workers known to have been exposed to dioxin, 54 who might have been exposed and 31 intended as a control group but drawn for convenience from senior office, laboratory and works staff, most of whom were being investigated for blood lipids at the time.

The 1977-78 survey showed blood-lipid disorders, with elevated concentrations of serum cholesterol and triglycerides, and depressed levels of high-density lipoprotein. The enzyme γ -glutamyl transpeptidase, diagnostic of liver function, was abnormally high in the exposed group. On the other hand, levels of two immunoglobulins (IgD and IgM) were depressed in the exposed group, suggesting impairment of immune as well as liver function.

The significance of these findings is far from clear, with the consensus among physicians in the field that further investigations are needed. They argue that the persistence of the phenomena uncovered by the earlier study should be confirmed — but also that further surveys should be more carefully controlled.

Dr George May, Coalite's company doctor, confirms that those workers included in the original survey have since been offered a second medical examination, but that "not everybody availed themselves of the opportunity". In the event, 29 of the original sample of 41 workers were re-examined, but measurements of immunoglobulin were not made.

There are no plans, however, for further investigations of the Coalite workers, according to Dr George Sorrie, deputy director of the Employment Medical Advisory Service, who says that his service has not been presented with any "realistic and useful" proposals for a long-term assessment of the health of the Coalite workers. The reassessment of the state of health of dioxin-exposed workers at

Dioxin a sarcoma risk

Reassessing clinical data is not the only reason for monitoring the health of workers exposed to dioxin. Recent letters in *The Lancet* (P A Honchor and W E Halperin, i, 268, 31 January 1980, and R R Cook, i, 618, 14 March 1980) highlight another reason.

The authors of the letters point out that in four industry-sponsored mortality studies of dioxin-exposed workers in the United States, 105 deaths were recorded. In four of these cases, death was caused by soft tissue sarcomas. These are rare forms of cancer and for most the aetiologies are unknown.

Soft tissue sarcomas have also been identified in another cohort of workers considered to have been exposed to dioxin. According to two Swedish epidemiologists, Dr Lennart Hardell and Anita Sandstrom of the Centre of Oncology at the University of Umea, exposure to the dioxin contaminated herbicide 2,4,5-T (2,4,5-trichlorophenoxyacetic acid) or chlorinated phenols 6 to 20 years previously could be responsible for the sixfold greater incidence of soft tissue sarcomas recorded in a survey of Swedish forestry workers (*British Journal of Cancer* 39, 711, 1979).

Alastair Hay

Coalite is thus far from complete. One clinician describes the re-examination of the 29 workers as a "totally useless piece of work" which cannot be expected to yield any meaningful results.

Coalite's managing director, Mr Charles Needham, has been approached by several clinicians prepared to carry out further investigations. Coalite's reluctance to agree to publication of the results from any future comprehensive medical screen of its workers has, however, forced at least one clinician to abandon his proposal.

Mr Needham, described by his secretary as solely responsible for press relations, has not been available for comment. One consequence of criticisms of the 1977-78 survey, and of Coalite's continued insistence on secrecy, is that the financial sponsor of the survey (an independent non-governmental source) will be unwilling to provide further support.

As the law stands, Coalite may have the last word. Hitherto, the company has been able to exploit an ambiguity in the 1974 Health and Safety at Work Act when deciding whether or not to release medical information on its workforce (see *Nature* 6 March 1980, p 2). The routine powers of the Health and Safety Executive to demand medical records of workers likely to have been affected by a product lapse when manufacture stops. Although Section 27 of the act allows the executive to disclose whatever information it requires (see *Nature* 1 May 1980, p 4), epidemiologists at the executive consider that a test case is required to clarify the law.

Alastair Hay

CORRESPONDENCE

Majority verdict

SIR — In a recent communication to *Nature* G G Simpson writes about a small group of cladists at the American Museum of Natural History that is not representative of the staff as a whole (*Nature* 26 March, p 286). In his view a majority of the staff rejects cladistic or Hennigian procedures. There are 32 professional zoological systematists on the staff here, not including emeriti or honorary staff appointments, and 26 of these use cladistic methods in their systematic work. The small, unrepresentative group referred to by Simpson therefore amounts to 81 per cent of the relevant staff. On the recommendation of three scientific departments, the Council of the Scientific Staff, and the Awards Committee, the trustees of the museum in 1975 gave Willi Hennig its Gold Medal for distinguished contributions to science.

J W ATZ, C J COLE, N ELDREDGE,
W K EMERSON, E S GAFFNEY,
B N HAUGH, L H HERMAN, E KIRSTEYER,
M C MCKENNA, J G MAISEY, C W MYERS,
G NELSON, N I PLATNICK, F H RINDGE,
D E ROSEN, J G ROZEN, R T SCHUH,
C L SMITH, I TATTERSALL, R H TEDFORD,
R WYODZINSKY AND, R G ZWEIFEL
*The American Museum of Natural History,
New York, USA*

'I was there'

SIR — Your account (*Nature* April 9, p 435) of a lecture I delivered "at the Rockefeller University in New York last week" raises in me the strong suspicion that your reporter did not hear what I said. This impression is strengthened by the fact that the lecture he reports was actually given in the Cornell University Medical Center. J L GOWANS
*Medical Research Council,
London W1, UK*

Dr Gowans' suspicion is unfounded. Our reporter indeed heard what was said and reported accurately that part of Dr Gowans' talk concerned with the applicability of the Rothchild principle to the financing of medical research in Britain. The Cornell University Medical Center is on the other side of York Avenue in Manhattan from the Rockefeller University, which had advertised the meeting at which Dr Gowans spoke. The two institutions run many joint seminar programmes, but are of course organisationally distinct. — Editor, *Nature*

Evolution's Waterloo

SIR — While no sensible person could disagree with your position on the "theory of evolution" (*Nature*, 12 March, p 75) I think your arguments are unnecessarily weakened by the sloppy use of words in multiple meanings, enabling your opponents to confuse the issue.

The phrase "theory of evolution" for example, is used in at least two quite different meanings. At one time it means "a record of events in the history of the Earth", at another it means "an explanation of the underlying causes of these events".

The term "metaphysical theory" is another example of this abuse. Popper notwithstanding, a metaphysical theory, *sensu stricto*, is a theory which cannot be verified (or

falsified) either now, or ever or anywhere, say the proposition "God exists". The record of historical events, say the proposition "Napoleon lost the battle of Waterloo" is not a metaphysical theory. Although the event cannot be rerun, it certainly can be verified in many different ways. Therefore the "theory of evolution" in its first meaning cannot be called metaphysical, technically or otherwise. An explanation of why Napoleon lost the battle of Waterloo, on the other hand, may or may not be metaphysical, according to whether it is based on verifiable facts or not. An explanation based on the intervention of an army of angels could be called metaphysical, one based on the superiority of the British and Prussian rifles may or may not be true, but it certainly is not metaphysical. The same reasoning holds for the "theory of evolution" in its second meaning.

*Administration de l'Industrie, S V VAECK
Brussels, Belgium*

SIR — As an American "cladist" and evolutionist concerned with the health of both disciplines and the public reaction to them, I feel that some discussion of the issues raised in *Nature* is warranted.

The issue of creation versus evolution entails two points, neither of which was mentioned in sufficient detail for the public readers of *Nature* properly to understand. First, the real controversy is over whether or not the phenomenon of evolution occurs. It does not concern whether Darwin or anyone else was correct about the mechanisms behind the phenomenon. To imply that a refutation of Darwinism or neo-Darwinism constitutes a refutation of evolution is like implying that a refutation of Newtonism is a refutation of gravitational attraction. Both conclusions would be misconceived. Second, the issue of creationism versus evolution certainly does not rest on whether Karl Popper's particular philosophy of science renders both creationist myths and Darwinian conjectures "untestable." There is a higher principle to resolve this issue.

Science is a discipline of open-minded inquiry which seeks to provide explanations about the world without invoking the supernatural. Creationism as practised by fundamentalist Christians is a close-minded doctrine built around an ultimate source of truth, the Bible. Creationism is dedicated to the proposition that the only explanations about "origins" are the supernatural explanations set forth in their ultimate source of truth. Thus "scientific" creationism, by any standard of rational scientific philosophy, is not science nor can it ever be science. Karl Popper is quite aware of this and has said so in *Conjectures and Refutations*.

Another issue concerns what is being called "cladistics." Scientists who fall under this label have diverse views about the proper relationship between "cladistics" and evolution. As one "out-and-out cladist" I cannot imagine something less interesting than a branching diagram (cladogram) without an evolutionary interpretation. Indeed, I cannot imagine beginning a study of biogeography without such an interpretation. I will even claim that an assumption of descent with modification is necessary to do such analyses. I base my claim on two points.

First, I have never seen a natural hierarchy

that was not due to historical descent in the larger sense of the phrase. This includes natural hierarchies in both the organic and inorganic worlds. Examples of natural hierarchies in the organic world are, of course, phylogenies of both organisms and languages, and the hierarchy of ontogeny of individual multicellular organisms. The only example of a natural hierarchy in the inorganic world that I know of is the hierarchy that continents display through descent from common "ancestral" continents. We observe the unfolding of ontogeny, we have detailed histories for the evolution of languages, we have a very good mechanism for the drifting of continents. Are we to think that the only other example of a natural hierarchy, the branching relationships between species, is different in kind from the other three?

Second, the only alternative to the evolutionary interpretation is the assumption of a logical creator who made each species (or group of species, or language, or continent) to look as if it evolved. As I wish to practise science rather than theological metaphysics, I make the assumption of evolution. If I wished to practise theological metaphysics, I would certainly not conceive of a logical creator who was, at the same time, tricky enough to make things look as if they evolved when they did not. I make the assumption of evolution without letting my hypotheses of evolutionary relationships (cladograms to my mind) be dictated by particular theories such as neo-Darwinism. If I let neo-Darwinism dictate *a priori* what relationships were possible then I would be no better than the fundamentalists who let the Bible dictate to them what is possible.

E O WILEY
*Museum of Natural History
and Department of Systematics & Ecology,
University of Kansas, USA*

No to sociobiology

SIR — The recent controversy on genes and racialism (*Nature* 22 January, 12, 19, 26 February) may have some relevance to caste system in India, since it has now assumed a great political significance and is creating social tensions. Although castes are manmade (all Hindus are born unequal!), and the system is nowhere near as old as the human race, yet attempts are often made to establish a biological basis of castism in line with the insect societies. Wilson (in *Sociobiology*) has refuted any genetical basis for caste system, yet willy-nilly, theoretical biology has become a convenient and manipulative tool to disdain or support such social practices.

The problem will become more complex and acute if attempts are continued to biologize every social attribute, be it racialism, castism or nazism. Biologization may be well-intentioned, and it may well prove that some of our beliefs are wrong and unscientific, yet by this process it may generate unnecessary, avoidable controversies and social conflicts. Surely, without biologizing, one can distinguish between races and racialism and between castes and castism. One may dispute or not give a damn if the human race is united or divided, but one can surely condemn racialism or castism without looking for biological evidence or defending the theory of kin selection. Let us keep sociology and biology apart.

B BANERJEE
Tea Research Association, Assam, India

NEWS AND VIEWS

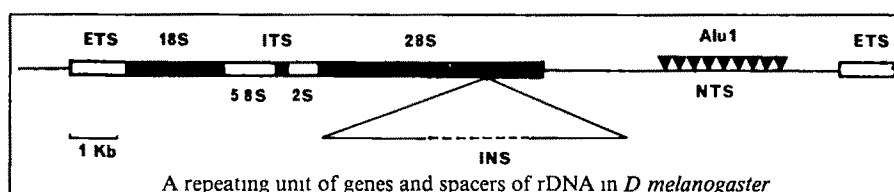
Springcleaning ribosomal DNA: a model for multigene evolution?

from Gabriel Dover and Enrico Coen

THE nuclear DNA of higher organisms is subject to such a diversity of sequence rearrangements that one might wonder how developmental or evolutionary stability is maintained. The variety of rearrangements, no doubt arising as errors in the regular processes of DNA replication and recombination, gives rise to seemingly randomly distributed families of multiple-copy sequences. A major unresolved question concerns the biological and evolutionary consequences of these mutational processes. One particular family of sequences that might provide the first insights into the rates at which different mutational mechanisms progress and of their possible phenotypic effects is composed of tandem head-to-tail arrays of ribosomal RNA genes (rDNA). From the data reported in this issue of *Nature*, p 749, by H Roija, J R Miller, L C Woods and D M Glover, and elsewhere²⁻⁵ it is now possible to observe in one family of repetitive sequences most, if not all, the phenomena usually depicted separately in different families.

The 18S, 5.8S and 28S RNA moieties of eukaryote ribosomes are co-transcribed from a unit of contiguous genes separated by internal and external transcribed spacers (ITS and ETS in the figure). Units are separated by non-transcribed spacers (NTS), in arrays that are located at the nucleolar organizers (NO). An exhaustive *variorum* account of the organization of rDNA and other multiple gene families has recently been compiled by Long and David⁶.

Since the pioneering studies of D D Brown, P K Wellauer, R H Reeder and I David into rDNA variation in species of *Xenopus*, it has become apparent that this family of repeats has been subjected to a new type of evolutionary force termed horizontal^{7,8} or concerted² evolution. This term is used to describe extensive sequence homology of NT spacers within a species



even though there is very little homology between spacers of related species. The mechanism responsible for concerted evolution of copies of a sequence within and between individuals should take into account the striking interspecific homology of the genes with which the spacers are regularly interspersed — for example^{3,9}, there is 93 per cent homology within 200 base pairs of the 28S gene of an insect³, an amphibian⁹ and a ciliate protozoan¹⁰ — and should result in rapid 'horizontal' spread of each new spacer variant to all rDNA units of all individuals of a species. Alternatively, it is conceivable, although unlikely, that both the rDNA arrays and populations go through drastic reductions in number followed by the establishment of new arrays from a few founder units. The recent analysis of sequence variation of rDNA in a variety of organisms, in particular in species of *Drosophila*, reveals several overlapping levels of evolutionary change that throw some light on the mechanism of concerted evolution and the complex progression of rDNA in general.

In *Drosophila melanogaster* there are two clusters of rDNA situated on the X and Y chromosomes. An interesting additional feature is that 60 per cent of the 28S genes of the X chromosome contain an insertion (INS-type I). Another non-homologous type II INS is present in 15 per cent of the 28S genes of both chromosomes. Although insertion sequences are by no means a universal feature of nuclear 28S genes (they are confined so far to the higher diptera of metazoan animals and to ciliate protozoa), they have been shown to contain sufficient residual interspecific homology to suggest that they are of ancient origin and arose before the divergence of the dipteran

suborder Cyclorrhapha into the calyptrates (such as the house-fly) and acalyptrates (such as *Drosophila*) approximately 65 Myr ago¹¹. Bearing in mind their age it is clear that if complete rDNA arrays are subjected to a concerted 'spring-clean', then the INS⁺ units are either selectively maintained or periodically re-inserted into newly amplified units. For the latter they would require a degree of mobility.

The data of Glover and Rae and their colleagues^{1,3} on the nucleotides surrounding INS elements in *Drosophila* elucidate this question of mobility. In *Drosophila virilis* there is a 14-base pair duplication of a 28S gene sequence that flanks the insert. This same sequence is present only once in INS⁻ units. The generation of a host sequence duplication (usually of 5–9 base pairs) upon insertion of an extraneous sequence is a common feature of mobile elements in *Drosophila* and yeast^{12,13}. In *D. melanogaster* the type I INS is flanked on its right by this same sequence and on the left there is a 9-base pair deletion in the 28S gene. The two inserts clearly slot into identical positions in the two species suggesting a requirement for some sequence homology between host and insert that is analogous to the integration of λ phages into *Escherichia coli*. These data underline the possibility of mobility of INS elements. It is unlikely that there are any fundamental differences between the units that receive INS and those that do not. For example, Long, Rebert and David¹⁴ find no differences between INS⁺ and INS⁻ units in 470 nucleotides around the transcription initiation sites and in the sensitivity to DNase I of the chromatin of genes, spacers and inserts. Is there, then, a limiting supply

Gabriel Dover and Enrico Coen are in the Department of Genetics, University of Cambridge

of INS elements to re-occupy the 60 per cent of X chromosome 28S genes with type I and the 15 per cent of X and Y 28S genes with type II, or are the numbers selectively constrained? Interestingly, extra copies of type I sequences have been located outside of the NOs in both *D melanogaster*⁴ and *D virilis*³ that are situated for the most part in the centromeric heterochromatin in *D melanogaster*⁵. However, type I INS units in the heterochromatin are each flanked by the same 28S gene segments that flank type I INS in the 28S gene¹, suggesting a transposition of INS with its flanking sequences from the NO to the heterochromatin and not vice versa.

The phenomenon of concerted rDNA evolution in *Drosophila* species has only recently come to light. Electron microscope observations on the extent of heteroduplex matching of interspecific NT spacers of sibling species of the *D melanogaster* species subgroup initially underscored, at this level of resolution, the conservation, not the dissimilarities, in spacers between the species¹⁵. However, our analysis of the distribution of restriction sites and sequence homologies of NTS, ITS and INS elements of the species nevertheless shows that all these components are divergent between species and yet have evolved in concert within each species to the extent that a complete replacement of one sequence variant by another has occurred in both X and Y NOs¹⁶. Similarly, the NTS sequences of two more distantly related species, *D melanogaster* and *D hydei*, have been found not to be the same¹⁷.

Although evidence for the precise mechanism responsible for within-species homogeneity is not available, there are observations that strongly hint at one possible cause. A variety of studies have shown that an heterogeneity in length and an internal repetition within NT spacers are common features of multiple-copy genes (see refs 7 and 18 for reviews). For example, the lengths of NT spacers both within and between sibling species of *D melanogaster* often differ by multiples of 250 base pairs¹⁶ which is the periodicity of the *AluI* repetition in the NTS of *D melanogaster* (see figure). It has been suggested^{7,8} that a process of unequal exchange within NT spacers, in which the inequality is of the same periodicity as the internal repetition, would generate the observed differences. Given such a mechanism plus the ability of sequences of DNA to transpose or be exchanged between chromosomes, it is possible to offer some explanation for the complete horizontal replacement, at several different levels, of the rDNA within and between individuals in one group of species.

First, within the NO of one chromosome it is possible to view the replacement of repeats, one variant by another, as an outcome of a continuous process of amplification and elimination of sequences (by unequal exchange) in a stochastic

manner. At the lowest level there would be the replacement of units of the internal repetition within the NTS, for example, in *D melanogaster* a replacement by a variant repeat (containing an *AluI* site) has taken place in an array of 250-base pair repeats (lacking the *AluI* site) that is probably common to the NT spacers of the sibling subgroup¹⁶. An analogous 250-base pair periodicity has been detected in the NTS of *D hydei*¹⁷. In addition, unequal exchange within the internal repetition would also give rise to the 250-base pair length heterogeneity described above. At a higher, level unequal exchange involving complete rDNA units would lead to the replacement of complete arrays of NT spacers with their variant internal repeats in any one NO.

The second step in the process of sequence encroachment would be the establishment of the same spacer variant in all NOs of a karyotype. N. Arnheim and his colleagues report the complete distribution of spacer variants to several NOs within any one species of primate², analogous to the distribution of *AluI* variants to the X and Y NOs of *D melanogaster*¹⁶. Such a mobility between chromosomes could be either a direct transposition or dependent on both meiotic (homologous) and mitotic (non-homologous) chromosome exchanges. Finally, and importantly, if the processes of unequal exchange and interchromosome spreading are rapid, then a variant will be driven throughout a population of sexually reproducing individuals the last stage of concerted evolution. The absence of much rDNA sequence polymorphism, as distinct from length heterogeneity, within species suggests that the mutation rate producing new variants might be low relative to the rates at which they displace and spread both within and between individuals.

Against such a background of periodic replacement of rDNA units, in *Drosophila*, the exclusion of specific INSs from the NOs of the Y chromosomes of *D melanogaster*⁶ and *D hydei*⁷ suggests that insertion is a separate and independent stage in the overall evolution of the rDNA clusters.

The role of selection in the evolution of rDNA is much harder to gauge. There is probably a selective constraint on the number of 28S genes that can accept an insertion given the small deletion¹ (Glover

op cit) and the absence of transcription of INS⁺ genes in somatic tissues of *D melanogaster*^{4,19} although the role of selection is ambiguous in strains of *Tetrahymena pigmentosa* which differ by the total absence or presence of insertions throughout the 28S genes¹¹. There might also be a selective constraint that restricts the length heterogeneity of NTSs to within the defined lengths observed between sibling species of *Drosophila*¹⁶ and within individually characterized X and Y NOs of *D melanogaster*^{4,16,20,21}. This might ensure the correct phasing of nucleosomes vis-à-vis the transcription initiation sites of the repeats. An interesting regularity in nucleosome phasing with this property has recently been revealed in multiple-copy families of histone and ribosome genes in *D melanogaster*²² and in 5S RNA arrays of *X laevis*²³.

One final consideration concerns the generality of the phenomenon of concerted evolution and repeat unit replacement in eukaryote genomes. It is known that repeated sequences are undergoing recurrent rounds of amplification and that on occasion a variant member becomes the fortuitous unit of amplification leading to the establishment of an homogeneous but variant array (segment) (see ref. 24 for references). Such polymorphic segments might represent a state of transition during replacement. Segmental variants and concerted replacement have been observed in tandem arrays of satellite DNAs, histones, immunoglobulins and globins, in addition to rDNA. Recently, evidence for concerted replacement has been found in a set of repeats that are dispersed throughout the genomes of different rodent species²⁵. This might result from a progressive 'gene conversion' of one interspersed repeat by another (rather than unequal exchange) along the lines described by Scherer and Davis in yeast²⁶.

The genomic consequences of sequence replacement, mobility and interchromosome spread, occurring simultaneously, for example, in the rDNA repeats, are clear: the rapid establishment of a group of genomes (species) that are uniquely identified by many shared homogeneous families of sequences. The biological consequences of the accidental fixation of multiple genomic dissimilarities between groups are endless. □

1. Rojja, H., Miller, J. R., Woods, L. C. & Glover, D. M. *Nature* **290**, 7-9 (1981).

2. Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O., and Zimmer, E. *et al Proc natn Acad Sci U S A* **77**, 7323 (1980).

3. Rae, P. M. M., Kohorn, B. D. & Wade, R. P. *Nucleic Acids Res* **8**, 3491 (1980).

4. Dawid, I. B., Lauth, M. & Wellauer, P. K. *Miami Winter Symp* **17**, 217 (1980).

5. Peacock, W. J., Appels, R., Endow, S. & Glover, D. *Genet Res Camb* **37**, 209 (1981).

6. Long, E. O. & Dawid, I. B. *A Rev Biochem* **49**, 727 (1980).

7. Brown, D. D., Wensink, P. C. & Jordan, E. *J molec Biol* **63**, 57 (1972).

8. Wellauer, P. K. & Reeder, R. H. *J molec Biol* **94**, 151 (1975).

9. Gourse, R. L. & Gerbi, S. A. *Nature* **283**, 693 (1980).

10. Wild, M. A. & Sommer, R. *Nature* **283**, 693 (1980).

11. Barnett, T. & Rae, P. M. M. *Cell* **16**, 763 (1979).

12. Dunsmuir, P., Brorin, W. J., Simon, M. A. & Rubin, G. M. *Cell* **21**, 575 (1980).

13. Farabaugh, P. J. & Fink, G. R. *Nature* **286**, 352 (1980).

14. Long, E. O., Rebbert, M. L. & Dawid, I. B. *Proc natn Acad Sci U S A* (in the press).

15. Tartof, K. D. *Cell* **17**, 607 (1979).

16. Coen, E. & Dover, G. A. (in preparation).

17. Renkawitz-Pohl, R., Glatzer, K. H. & Kunz, W. *Nucleic Acids Res* **8**, 4593 (1981).

18. Fedoroff, N. V. *Cell* **16**, 697 (1979).

19. Jolly, D. J. & Thomas, C. A. *Nucleic Acids Res* **8**, 67 (1980).

20. Indik, Z. K. & Tartof, K. D. *Nature* **284**, 477 (1980).

21. Endow, S. A. *Cell* **22**, 149 (1980).

22. Samal, B., Worcel, A., Louis, C. & Schedl, P. *Cell* **23**, 401 (1981).

23. Gottesfeld, J. M. & Bloomer, L. S. *Cell* **21**, 751 (1980).

24. Dover, G. A. *Nature* **285**, 618 (1980).

25. Brown, S. D. M. & Dover, G. A. *J molec Biol* (submitted).

26. Scherer, S. & Davis, R. W. *Science* **209**, 1380 (1980).

Continental margin processes

from David G Roberts

OVER the last decade research on continental margins has been greatly stimulated by the need to assess the hydrocarbon potential of the largely unexplored margin areas and to develop a thorough geological and geophysical understanding of the transition across the fundamental crustal boundaries defined by passive and active margins. The recent Hedberg Conference* was the third to review continental margins since the original Penrose Conference of 1972 (see *The Geology of Continental Margins* Springer Verlag, 1975).

Results obtained from seismic exploration allied to the International Phase of Ocean Drilling (IPOD) were reviewed by A W Bally (Shell Oil, USA). Specific problems of passive margins revealed by the different structural styles of the north Biscay and eastern North American margins include the mechanism and amount of stretching during rifting and its effect on the subsequent history and thermal subsidence. In the case of active shear and convergent margins the structural implications of brittle-ductile boundaries in the lower crust have yet to be fully understood as has the role of crustal extension in an active margin setting. Bally urged full documentation of the palaeontological data base used to erect the global sea-level curves of Vail *et al* (*Mem Am Ass petrol Geol* 26; 44, 1981) and suggested changes in the subsidence rate of margins due to global sea-level variations as an alternative mechanism.

Convergent margins were reviewed by Von Huene (USGS, Menlo Park). IPOD drilling and associated seismic surveys have revealed evidence of uplift and tectonic erosion that have radical implications for mechanisms of thrusting and subduction. Subduction erosion or underplating has been proposed to account for tectonic erosion off Central America while uplift on the Japanese active margin may result from development of a normal, that is, more elevated geothermal gradient during periods of inactive subduction. Recession of the continental edge by as much as 60 km from the trenches may now be accounted for by dewatering and temperature-controlled 'clay' mineral transformations with associated faulting in response to weak shear stress conditions in the accreting sediments.

Other active margins discussed included the Nankai Trough off Japan (Aoki, Japan Petroleum) and the Makran margin where uplift related to convergence is about

1.5–2 km per million years (White, University of Cambridge). Examples from the Caribbean included the margins of Venezuela and Hispaniola where studies of seismic reflection profiles offshore have been integrated with a field study of the exposed parts of the active margin onshore (Biju-Duval, IFP, France). Along the Indonesian Arc (Curry, Scripps Institution) variations in structural style are reflected in the development of strike-slip faults and spreading back arc basins in turn related to the direction of convergence. Episodicity in the geological history of the exposed parts of the Arc may be more apparent than real and is perhaps caused by changes in plate convergence. Uplift and subsidence of part of the Aleutian Arc were also documented by Scholl (USGS, Menlo Park).

The Caledonian–Appalachian orogenic belt was the subject of several interesting interpretative comparisons with the marine results. Leggett (Imperial College, London) demonstrated the particularly striking resemblance between the Southern Uplands of Scotland, the Makran and the Middle American Trench while McKerrow (University of Oxford) reviewed the history of the northwestern edge of the Iapetus Ocean. Evidence of very large-scale tectonic transport along convergent margins was documented from California and the Appalachian belt. 'Suspect' terranes in California show 30–40° of rotation and transport of Upper Cretaceous rocks by as much as 3,000 km (Howell, USGS, Menlo Park). COCORP results in Appalachia also demonstrate thrusting of several hundred kilometres along the Blue Ridge thrust and have wider implications in interpreting autochthonous terranes in the Caledonide–Appalachian belt.

The session on passive margin evolution included reports from all over the world. Grachev (Institute Physics of the Earth, USSR) reviewed the Moma rift — the onshore continuation in Siberia of the spreading Gakkel Ridge of the Arctic Ocean. Nunns (Durham, UK) discussed the evolution of the Norwegian Sea and the Jan Mayen microcontinent. The structure of the entire Southwest African margin from the Walvis Ridge to Capetown was presented by Gerard (SOEKOR, South Africa) using seismic reflection and commercial results. This margin, like the Voring Plateau, East Greenland and some other margins, is characterized by a thick sequence of oceanward dipping reflectors close to the continent–ocean boundary. Drilling suggests that these reflections may be basalts. Association of supposedly oceanic magnetic anomalies with

continental crust on this margin questions their use in defining the continent–ocean transition. The Congo–Cabinda area further north is a classical area of rift tectonics where the entire transition from the pre-rift phase through rifting, initiated in the late Jurassic, to the Albian post-rift and the important role of halokines is well documented from seismic and drilling results (Cochran *et al* Gulf Oil, USA).

Off eastern North America, improved seismic data and the results of recent IPOD legs have shed new light on the evolution of the margins. In the Straits of Florida (Buller and Schwab, UTMSI) drilling on partly exposed tilted blocks has cored gneisses and phyllites. Careful analysis of refraction (Ibrahim and Uchupi, WHOI) and gravity and magnetic data (Hall, Gulf Oil) from the Gulf of Mexico suggest rifting and spreading formed a rhomboidal oceanic in early Jurassic time. Independent evidence for early Jurassic opening of the western North Atlantic has been given by the results of DSDP hole H534 which cored basement of Callovian age east of the Blake Escarpment (Sheridan, Gradstein *et al*, University of Delaware). In the nearby Carolina Trough, the East Coast magnetic anomaly is interpreted as an edge effect associated with the continent–ocean boundary by Hutchinson *et al* (USGS–WHOI) who also find no evidence of a basement ridge along the boundary. They propose a model of rapid stretching and subsidence to account for a zone of attenuated continental crust landwards of the boundary. Present growth faulting on this margin is apparently related to tectonically controlled halokinesis (Dillon, USGS, WHOI).

Models of stretching and subsidence provided one of the main points of formal and informal discussion. McKenzie (*Earth planet Sci Lett* 40, 25, 1978) originally proposed a simple model of lithospheric stretching and applications of this model were discussed by several speakers. Sobolev and Artemjev (Institute Physics of the Earth, USSR) discussed rifting and stretching in terms of phase transitions within the mantle. Variations in the width of passive margins may be related to eclogitization. Keen *et al* (Bedford Institute, Canada) analysed the development of sedimentary basins off North America in terms of the McKenzie model, models of uniform extension with melt segregation and of depth-dependent extension requiring rheological zonation within the lithosphere. They compared calculated with observed gravity, stratigraphy, mantle depth and thermal history data and concluded that the depth-dependent extension model gave the best fit. A rather different view was proposed by Sibuet and Le Pichon whose

*The Hedberg Conference was held from January 12 to 16, 1981 and was convened to honour Hollis D. Hedberg, Professor of Geology at Princeton and former chief geologist of Gulf Oil. The meeting was sponsored by the American Association of Petroleum Geologists.

David G Roberts is a Principal Scientific Officer at the Institute of Oceanographic Sciences, UK.

calculations of the crustal extension in Biscay were consistent with the uniform stretching model of McKenzie and could be applied to other areas. Chenet *et al* (IFP, Paris, France) also calculated the extension for the north margin of Biscay and Galicia Bank and calculated that a two-layer model best fitted the data although a mass balance problem remained in the lower crust. Slater *et al* (MIT) reviewed subsidence of the Georges Bank in terms of the McKenzie model and discussed possible hydrocarbon potential in terms of the thermal history. A two-layer finite element crustal model used to model the formation of outer highs by Vierbuchen and George (Exon, Houston) included variables such as the temperature at the base of the lithosphere and time-dependent creep. Watts (Lamont-Doherty, New York, USA) reviewed the roles of eustasy, flexural strength, diagenesis, stretching and cooling in margin subsidence, concluding that at least off North America the paucity of well data and the unknown thickness of syn-rift sediments do not easily allow test of McKenzie's model.

Sedimentary processes in many continental margin environments were documented using a variety of techniques. Coumes (SNPA, Pau, France) examined the depositional history of the Cap Ferret Fan, south-west Biscay using Seabeam and high-resolution seismic data linked to cores. A wider view of continental margin sedimentary processes was given by Roberts (IOS, UK) from the results of long-range sonar (GLORIA) surveys. Complete sonar mosaics of entire continental margins are giving a needed overview of the role of downslope and slope sediment movement in determining facies distribution. The role of slope parallel and downslope sediment transport in the development of the lower continental rise off the eastern USA was documented using seismic profiles (Laine and Tucholke, WHOI). A detailed study of the McKenzie River delta was presented by Cole (Gulf, Canada) who recognized five deltaic cycles in the delta's development. Bouma (USGS, Corpus Christi) discussed the influence of halokinesis on upper slope sedimentation in the Gulf of Mexico.

The maturity and characteristics of organic matter on continental margins was appropriately discussed in the final session. Arthur (USGS, Denver) discussed the role of Mesozoic-Cenozoic anoxic events determining the production and presentation of organic matter along the slope and rise. Accumulations are episodic and themselves variable in organic matter and type. A detailed geochemical study of Cretaceous organic matter by Katz (Texaco, USA) showed differences between the North and South Atlantic Oceans. The North Atlantic organic matter is predominantly of terrestrial origin and is likely to yield gas compared with the oil potential of the marine organic matter in the South Atlantic Ocean. □

Solar physics at Taos

from R. Rosner

THE second major presentation of results from the Solar Maximum Mission (SMM) satellite, which recently met an untimely failure in its pointing system after approximately six months of active operation, provided the focal point at the winter meeting of the American Astronomical Society's Solar Physics Division*. This gave the solar community an excellent overview of both the current status and the likely scientific impact of the SMM venture. Several other major results not associated with the SMM programme were also discussed.

One of the first new results reported concerns periodic fluctuations in hard (≥ 20 keV) X-ray emission from solar flares. Such fluctuations have been observed by both the SMM Hard X-Ray Burst Spectrometer (HXRBS), reported by Kiplinger (Goddard Space Flight Center), and an X-ray spectrometer on board the ISEE-3 spacecraft (Kane, UC Berkeley), who also reported corresponding quasi-periodic variations in microwave radio bursts. The typical pulse period is ~ 8 s, on a coherence time scale of the order of several minutes. Several explanations were offered for these observations, one plausible model proposed by Morrison and Ionson (Goddard) considers the flaring 'loop' structure as a lumped resonant circuit whose solutions on short time scales successfully predict oscillatory behaviour with periods of the order of seconds.

Detailed flare plasma diagnostics have been considerably advanced by SMM observations. High-quality observational evidence for both systematic flows and turbulent velocities were reported by the Ultraviolet Spectrometer and Polarimeter (UVSP) and X-Ray Polychromator (XRP) experiment teams, for example, turbulent velocities (as measured by broadening of lines due to high ionization states of calcium and iron) were shown to attain values up to ~ 100 km s⁻¹ during flares. Data from these two instruments and from the HXRBS have been assembled by Frost, Orwig, Poland and Shine (Goddard) who reported evidence that considerable hard X-ray flare emission originates from the vicinity of flare loop footpoints. These data, together with high-resolution Very Large Array (VLA) observations (Kundu, University of Maryland, and collaborators), strongly support the view that a non-thermal component is present in the electron velocity distribution and is responsible for

the footpoint hard X-ray emission. Yet further evidence for non-thermal processes comes from the Gamma-Ray Spectrometer (GRS) Forrest (University of New Hampshire), Ramaty (Goddard) and collaborators showed spectra which clearly indicated line emission at, for example, 2.2 MeV (attributable to neutron capture by protons) and 0.51 MeV (corresponding to the well known e^-e^+ annihilation line), and demonstrated convincingly that both ions and electrons must be accelerated impulsively (and episodically) during a flare. On similar time scales, the detailed temporal variation of the 2.2-MeV line flux was shown to follow the hard X-ray flux variation and to correspond to the behaviour expected from neutron slowdown and eventual capture. In a review of the high-energy experimental flare results from SMM, P. Hoyng (Utrecht) argued that the preponderance of evidence now points towards the 'non-thermal' flare models (although, as pointed out by A. G. Emslie (Stanford), spatial characteristics of hard X-ray emission which may be attributed to a non-thermal model can also be predicted by some 'thermal' models).

The considerable flare emphasis of SMM may lead one to forget that one instrument aboard SMM was especially designed to observe phenomena on longer time scales — ACRIM, a cluster of three active cavity radiometers designed to monitor the possible variability of the solar bolometric luminosity. This important, but controversial subject was reviewed by S. Sofia (Goddard), and Hudson (UC San Diego), Woodard (UC San Diego) and Willson (JPL) reported luminosity fluctuations as large as ~ 0.2 per cent (probably associated with the passage of sunspot groups across the solar disk). ACRIM should help to answer some of the outstanding questions in this area, for example, it can now be shown observationally that the well known sunspot energy 'deficit' is not spatially redistributed in the vicinity of a spot on short time scales, suggesting that transient energy storage does indeed occur in the solar convection zone.

There is, of course, an enormous amount of solar physics conducted outside the SMM realm. J. Kohl (Harvard-Smithsonian CfA) and collaborators are using resonantly scattered Ly α photons as a diagnostic tool for coronal plasma above 1.5 solar radii. They have had two successful rocket flights, and G. Withbroe (Harvard-Smithsonian CfA) reported one of the first major results — that proton and

*The meeting of the Solar Physics Division of the American Astronomical Society was held 7–10 January 1981 at Taos, New Mexico.

R. Rosner is at the Harvard-Smithsonian Center for Astrophysics.

electron temperatures rapidly decouple above ~ 2 solar radii. The relative spatial constancy of the electron temperature, in contrast to the rapidly declining proton temperature with increasing radius, implies that the coronal heating process primarily couples to electrons, and is consistent with no significant proton heating (other than via e^- - p coupling) above ~ 2 solar radii. IMP satellite data (Borini, Stanford, and Gosling, Los Alamos) included the first definitive identification of (at least one of) the coronal sources of the low-speed solar wind, that is, the 'belt' of coronal streamers overlying photospheric neutral lines, and anomalously low He III abundances in the high plasma density wind regions associated with sector boundaries and (when mapped back to the Sun) coronal streamers. It is still debatable

whether yet other coronal source regions of the low-speed solar wind exist, these IMP results (particularly the He abundance anomaly), however, provide an important constraint for solar wind theorists.

Although experimenters did rather steal the show, theorists too were present at the meeting. Amongst the interesting theoretical results reported were the hydrodynamic calculations of the chromosphere, transition region and corona (McClymont and Canfield, UC San Diego, Boris and Mariska, NRL, and Pallavicini, Arcetri Observatory) which promise to resolve the controversy on the structure and stability of the 'confined' solar atmosphere, and convection zone calculations (Hurlburt and Toomre, JILA) which may help understanding of convection in the fully compressible limit. $\square$

neutron electric dipole moment, where such very small neutron spin rotations must be measured.

The present experiment, an Ispra/Harvard/Rutherford/Sussex collaboration, may be viewed as a neutron version of the standard crossed-polarizer configuration used in optics. Polarized neutrons pass from a polarizer through the tin samples, which are about 5 cm long, to a symmetric analyser, which yields neutron counts depending on the degree of spin rotation. Care must be taken to understand and eliminate magnetic field effects which could give spurious rotations by acting on the neutron magnetic moment, and various control and null tests were performed.

In classical fashion, the effect first showed up not in the material being studied (^{124}Sn) but in a control sample (natural tin) from which a large effect was finally located in ^{117}Sn .

Results for the three materials are ^{124}Sn , $(0.48 \pm 1.49) \times 10^{-6} \text{ rad cm}^{-1}$, ^{117}Sn , $(36.7 \pm 2.7) \times 10^{-6} \text{ rad cm}^{-1}$, $\text{Sn}(\text{natural})$, $(4.95 \pm 0.93) \times 10^{-6} \text{ rad cm}^{-1}$. Since natural tin contains only 7.6 per cent ^{117}Sn , there would appear to be a residual effect of $(2.16 \pm 0.95) \times 10^{-6} \text{ rad cm}^{-1}$ in natural tin not explained by its ^{117}Sn content. It should be noted that large parity-violating effects have shown up in ^{117}Sn before, in neutron radiative capture⁶.

In the context of the present picture of weak interactions, estimation of the size of the spin rotation effect would seem to be straightforward⁵. The neutron-atom weak interaction contains a term $\sigma \cdot p$, proportional to the helicity of the neutron, which creates a different index of refraction for each of the two helicity states of the neutron in matter. This in turn induces a rotation of the spin component perpendicular to the motion. The rotation in radians, Φ , for a neutron travelling a distance Z in a simple Born approximation calculation is

$$\Phi = \sqrt{2} G W p Z$$

where G is the Fermi weak interaction constant, ρ the number density of atoms and W a 'weak charge' for the atom. W is expected to be on the order of the number of particles in the atom, about 70 for example in the currently accepted electroweak theory of weak interactions. Among the

Neutron optical activity

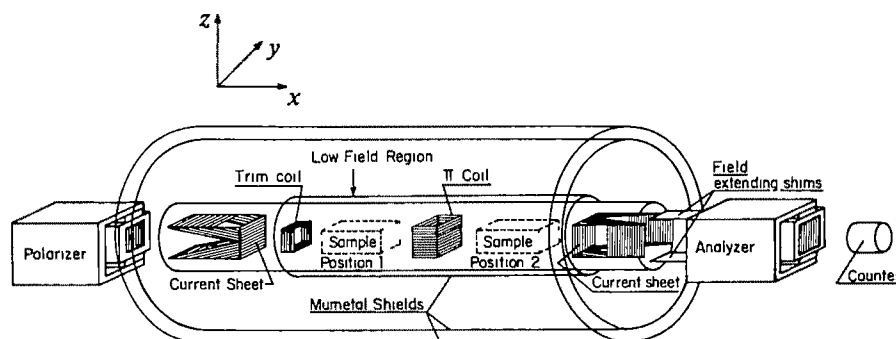
from Leo Stodolsky

THE PHENOMENON of optical activity, in which a substance rotates the plane of polarization of light passing through it, is a familiar one. It is a common property of organic chemicals and is often used, for example, to measure their concentrations. Recently, a group working at the ILL (Institut Laue-Langevin) in Grenoble has reported¹ the first observation of the equivalent phenomenon for matter waves. Furthermore, the size of effect appears, in at least one case, to greatly exceed simple theoretical expectations. In an experiment using cold neutrons from the ILL research reactor (see the figure), Forte and his colleagues have seen the rotation of the spin of the neutron around its line of flight as it propagates in solid tin. A neutron with its spin direction perpendicular to its motion is a linear combination of two helicity (or circular polarization) states, just like a linearly polarized photon, therefore such a neutron beam is exactly analogous to plane-polarized light.

The neutron spin rotation observed differs from the optical activity of chemicals in one fundamental respect, however. In both effects a rotation (of the spin or plane of polarization) is coupled with a direction (line of flight) and demonstrates a 'handedness' or parity sense associated with the material. In the case of ordinary optical activity this handedness is traceable to the physical structure of the molecules of the material, which are in a configuration of definite handedness (stereoisomerism). A sample of tin, however, contains no such handed structures (The possibility of a neutron spin rotation created by handed materials,

through nuclear forces², or through magnetic interactions³, has been considered, but the predictions are quite small.) Thus the parity sense manifested by the neutron beam must be traced ultimately to a handedness in the interaction forces between the neutron and the constituents of the material — that is, to the phenomenon of parity violation in the fundamental interactions. Thus the small (10^{-6} rad) but definite rotation measured is a striking manifestation of a parity sense imparted to bulk matter by parity-violating forces.

The possibility of such a 'neutron activity' was, in fact, first suggested⁴ by Michel in 1964 as a way of seeing the parity-violating, weak interaction forces between the neutron and the atomic nucleus. The interest of the effect in connection with the question of 'neutral weak currents' and the possibility of observing the neutron-electron weak interaction was recognized seven years ago⁵ and brought to the attention of Norman Ramsey (Harvard University), who has been involved in a long series of experiments searching for a



Schematic view of experimental apparatus. The neutron beam polarization is along z and momentum along x .

Leo Stodolsky is at the Max-Planck Institute for Physics and Astrophysics, Munich

interesting features of the formula are that it is proportional simply to G and that, in contrast to ordinary optical activity, which must vanish at low frequency, it is energy independent. The observation of the effect with such low-energy (millivolt) neutrons appears to verify this last feature.

The size of the observed effect is surprising, however. The above theoretical formula for Φ , as well as more sophisticated calculations (for discussion and references, see ref. 7), lead to an estimate of about 10^{-8} rad cm^{-1} . This is 3,000 times smaller than that observed for ^{117}Sn .

The only explanation so far put forward to explain such a large discrepancy involves a very low-energy neutron (P-wave) resonance. In fact, the original motivation⁸ for studying ^{124}Sn — where, ironically, no effect at the 10^6 rad cm^{-1} level was found — was the suggestion of a resonant enhancement for the effect. In ^{117}Sn a weak resonance, which could be P wave, has been reported (1.2 eV). Even with the existence of such a resonance, however, it seems difficult to get an effect of the observed magnitude. This has prompted at least the present author to wonder whether there may not be some surprises lurking in the still not very well understood sector of the fundamental interactions concerning the non-leptonic weak interactions, particularly the weak interactions between nucleons⁸. The imperative question here is whether the large rotation is indeed some kind of special nuclear physics effect particular to ^{117}Sn and perhaps a few other rare nuclides, or whether there is something unexpected generally residing in the nucleon-nucleon force. In this respect the residual rotation for natural tin, beyond that expected from its ^{117}Sn content, is intriguing. If true, it indicates a large rotation on some other isotope also, and makes it much less likely that what is happening with ^{117}Sn is a freak occurrence. These questions can be clarified by measurements on a variety of nuclei, particularly the lighter ones, where the resonances are widely spaced and well known. Should the effect consistently turn out to be too large, it could point to the need to revise our understanding of the weak interactions on a basic level.

Finally, another interesting fact has emerged from this experiment. Motivated by the large rotation on ^{117}Sn , Forte and his co-workers looked for a dependence of the cross-section on the helicity of the neutron, that is, a difference in the rate of

neutrons transmitted through the material according to whether the spin is parallel or antiparallel to the direction of motion. This is also manifestly parity violating, and quantum mechanically is governed by (the imaginary part of) the same amplitude as the spin rotation. Here again a large effect is found. Although the authors say the result is too preliminary to be quoted as an actual cross-section, it is stated as being at about 9×10^{-6} of the total cross-section. The magnitude of this may be appreciated by saying that it corresponds to a cross-section at the tens or hundreds of μb level, a number more typical of say, electromagnetic than weak interactions, and hence appears 'big'.

This observation is something of an embarrassment to the theoretical workers in the subject. Although we had considered

the polarization dependence of the cross-section, the possibility that it could be finite, let alone 'big', at threshold was never considered since it would appear to involve inevitably a P-wave suppression factor. A finite parity-violating cross-section at zero energy is, in fact, possible, however, when an exothermic channel such as radiative capture or perhaps fission is present. This leads to an amusing and complex situation where, for example, the dominant process for the P-wave scattering is parity violating. Furthermore, because of the spin structure, some interesting tests of time reversal invariance should be possible in this novel system.

Although it is not at all clear that this situation is fully understood, it does seem clear that the door has been opened here to a new set of very interesting phenomena. □

Two DNA-binding proteins

from David Davies

THE last few years have seen dramatic developments in our view of DNA as a dynamic molecule, both in structure and function. Most of these properties can be attributed to the proteins that interact with DNA to produce replication, transcription, translation, transposition and processing. There are polymerases, synthetases, untwisting proteins, twisting proteins, histones, repressors, activators, but until recently very little was known about their three-dimensional structure or how, precisely, they interact with nucleic acid. This has now begun to change with the X-ray diffraction studies reported in this issue of *Nature* on two DNA-binding proteins, the cro repressor of bacteriophage λ (p 754) and the *E. coli* catabolite gene activating protein (CAP) (p 744). Although neither crystal structure actually involves the protein-DNA complex, the protein structures in both cases immediately suggest how the complex might form. The remarkable conclusion is made that while cro binds to right-handed B-DNA, CAP binds to left-handed B-DNA.

Cro is a small protein (molecular weight 7,351) that, together with the λ repressor protein, acts in the control of the lytic pathway of the bacteriophage³. The structure of the monomer at 2.8 Å resolution, as determined by Anderson *et al.*¹, is very simple, consisting of three extended segments forming an antiparallel β sheet, together with three helices. An

extended C-terminal tail interacts strongly with the corresponding tail of another monomer to form a dimer. This dimer associates in the crystal with another such dimer to form a tetramer, although in solution it is not known whether cro is a dimer or tetramer. However it is known that it binds to three sites on λ DNA, each of which exhibits some two-fold (palindromic) symmetry suggesting that the interacting form of the protein will be at least a dimer with two-fold symmetry.

Insight into the structure of the cro-DNA complex has come from a combination of genetics, chemistry and model building. Cro binds to a 17-nucleotide length of the DNA and prevents alkylation of the phosphates and methylation of the guanines at a site in the major groove. Cro appears to bind principally in the major groove of the DNA since methylation of adenine in the minor groove is not prevented by binding.

What parts of the cro structure could be involved in binding to DNA? One very suggestive feature is the helix from Gln 27 to Ala 36, which protrudes from the surface of the protein. The corresponding helix in the other subunit of the dimer is 34 Å away and each helix is inclined at 32° to the line joining their midpoints. These are the dimensions of the grooves on B-DNA and model building indicates that the helices can be positioned to fit convincingly into the major groove of right-handed B-DNA. Between the helices, some residues of the C-terminal segment form a pair of antiparallel β -sheet strands parallel to the minor groove which may interact

- 1 Forte, M. *et al.* *Phys. Rev. Lett.* **45**, 2088 (1980).
- 2 Kabir, P. K., Karl, G. & Obryk, E. *Phys. Rev. D* **10**, 1471 (1970).
- 3 Harris, R. A. & Stodolsky, L. *J. chem. Phys.* **70**, 2789 (1979).
- 4 Michel, F. C. *Phys. Rev.* **133**, 3329 (1964).
- 5 Stodolsky, L. *Phys. Lett.* **50B**, 352 (1974).
- 6 Benkoulou, H. *et al.* *Phys. Lett.* **71B**, 287 (1977).
- 7 Karl, G. in *Fundamental Physics with Reactor Neutrons and Neutrons* (ed. von Egidy, T.) (Institute of Physics, London, 1977).
- 8 Forte, M. in *Fundamental Physics with Reactor Neutrons and Neutrons* (ed. von Egidy, T.) (Institute of Physics, London, 1977).
- 9 Stodolsky, L. *Phys. Lett.* **96B**, 127 (1980).

David Davies is in the National Institutes of Health, Bethesda, Maryland

with it

Another regulatory protein, CAP, forms a complex with cyclic AMP and binds to specific DNA sites thereby altering the rate of transcription. The structure of the complex CAP-cAMP has been determined by McKay and Steitz at 2.9 Å resolution. The molecule is a dimer. Each subunit contains two domains joined by a 4–5 amino acid hinge region. There is a cleft between the two domains and in one subunit this cleft is open while in the other it appears closed, implying flexibility of the hinge analogous to that observed in immunoglobulins. The larger N-terminal domain interacts with the corresponding domain of another subunit to form a dimer in which there is local two-fold symmetry between the domains. There appears to be no interaction between the two smaller C-terminal domains.

As in the cro protein, the obvious candidates for binding to the DNA are two protruding helices, one from each of the C-terminal domains. However, unlike cro, these helices are oriented in such a way that they cannot fit into the grooves of right-handed B-DNA but could bind to the major groove of left-handed B-DNA. Such a model of the complex can account for the genetic and chemical data in a manner that could not be done with any model involving right-handed DNA.

Supporters of left-handed DNA recently received a boost with the discovery of the 'Z' structure for the alternating copolymer (dG-dC)⁴. However the [CAP-cAMP-DNA] complex proposed by McKay and Steitz involves left-handed B-DNA and not Z-DNA which has quite different helical dimensions. In 1979 Crick *et al.*⁵ examined the evidence on the handedness of DNA and concluded that there was little doubt that it was right-handed. Strong support for this conclusion has come from the recent dodecanucleotide crystal structure of Wing *et al.*⁶ who showed that the molecule is a right-handed B double helix. A modelling study by Gupta *et al.*⁷ has demonstrated that a left-handed model can be constructed for B-DNA, but experimental evidence

strongly favours the right-handed model.

McKay and Steitz therefore hypothesize that the handedness of the DNA at the binding site must change when it complexes with CAP. In view of the apparent flexibility of the protein hinge and the loose attachment of the DNA-binding domain to the major domain of CAP, this change of hand is a surprising accomplishment. McKay and Steitz have examined the possibility of rearranging the relative locations of the helices, for example by rotation about the hinge, but conclude that no simple rearrangement will provide a configuration complementary to right-handed B-DNA. Proof of this model will have to come from unwinding experiments on closed circular DNA, or from the crystallization of the CAP protein complexed to a small DNA fragment of appropriate sequence.

Given the complexity of protein structures, it is interesting that both the cro and CAP proteins have structures that immediately suggest that their principal means of interaction with the DNA is through α -helices sticking in the major grooves. The only crystallographic analysis of a nucleic acid-protein complex has been

that of protamine-tRNA⁸ which showed an α -helical piece of protamine lying in the minor groove of the tRNA. However, there have been other modelling proposals involving β -structure^{9–11} and the cro protein may also utilize the binding of an anti-parallel pair of strands in the minor groove of the DNA. Analysis of further DNA-binding proteins may reveal a more complex pattern of interactions although it should be remembered that DNA is a groovy molecule that might favour the binding of the ordered segments.

Although a model of DNA sequence recognition is not yet revealed by these studies, it is interesting that both proteins are at least dimers and can thus impose some 2-fold symmetry on the DNA binding site. Further information can be expected to come from a refinement and higher-resolution analysis of these and other structures. Also, although the three-dimensional structures of cro and CAP are quite acceptable, being based on experimental analysis with a proven technique, the predictions about the complexes are less secure and illustrate the need for crystal structure determinations directly of complexes of protein and DNA.

Laminin and epithelial cell attachment

from Brigid Hogan

THE past few years have seen an increasing interest in epithelial cells and the factors responsible for maintaining their organization into sheets and glands. The subject is important because most human tumours arise within epithelia, and metastasis involves cells breaking away from the territorial boundaries of these tissues and invading the surrounding stroma. Also important is the question of what distinguishes this pathological invasion by malignant carcinoma cells from normal migrations seen during the embryonic development of organs like the breast, when cords of cells burrow from the surface into the underlying mesenchyme to form complex branching ducts and tubes.

Such questions have focused attention on the physical boundary between epithelium and the stroma. In nearly all cases this is an extracellular basement membrane, or basal lamina, made up of proteoglycans, Type IV procollagen, fibronectin, laminin and other glycoproteins. Apart from restraining the wayward migration of epithelial cells by promoting their attachment and

alignment, roles suggested for the basal lamina include a means of laying down specific landmarks for cell guidance, acting as a filter for molecules of a particular size or charge, and the induction of epithelial cell differentiation. Given all this speculation, it is not surprising that the recent identification of the glycoprotein laminin¹ as a major constituent of basement membranes, both in the adult² and the embryo^{3–5}, has generated a whole spate of experiments on the effect of the purified protein on epithelial cells in culture.

Conveniently for the biochemist, laminin is made in large amounts by a number of mouse tumours, including the EHS 'sarcoma' (see ref 1) and several derived from an embryonic tissue known as parietal endoderm⁶. Laminin isolated from the matrix of these tumours consists of a very high molecular weight (approximately 900K), disulphide-bonded complex of three polypeptide chains, a larger one of about 450K and two smaller ones of very similar molecular weight (about 230K). Antibodies to laminin also immunoprecipitate a fourth polypeptide of molecular weight 150K, which is made in much larger amounts by normal rather than tumorigenic endoderm cells⁵, but the role of this laminin-related protein is not yet clear.

Brigid Hogan is in the Mammalian Development Laboratory, Imperial Cancer Research Fund Laboratories, Mill Hill, London

- 1 Anderson, W. F., Ohlendorf, D. H., Takeda, Y. and Matthews, B. W. *Nature* **290**, 754 (1981).
- 2 McKay, D. B. and Steitz, T. A. *Nature* **290**, 744 (1981).
- 3 Ptashne, M., Jeffrey, A., Johnson, A. D., Maurer, R., Meyer, B. J., Pabo, C. O., Roberts, T. M. and Sauer, R. T. *Cell* **19**, 1 (1980).
- 4 Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G. and Rich, A. *Nature* **282**, 680 (1979).
- 5 Crick, F. H. C., Wang, J. C. and Bauer, W. R. *J. molec. Biol.* **129**, 449 (1979).
- 6 Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R. E. *Nature* **287**, 755 (1980).
- 7 Gupta, G., Bansal, M. and Sasisekharan, V. *Int. J. Biol. Macromols.* **2**, 368 (1980).
- 8 Warrant, R. W. and Kim, S.-H. *Nature* **271**, 130 (1978).
- 9 Carter, C. and Kraut, J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 283 (1974).
- 10 Church, G. M., Sussman, J. L. and Kim, S.-H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1458–1462 (1977).
- 11 Blake, C. C. F. and Oatley, S. J. *Nature* **268**, 115 (1977).

There is no doubt, however, that laminin is completely distinct from fibronectin, that other widespread component of basement membranes. In contrast to fibronectin, laminin is only made by epithelial cells and a few others including myotubes and neuroblastoma⁷ and not by fibroblasts. Although it is secreted into the medium of cultured cells, it does not by itself form an extensive fibrous network on both the upper and lower surfaces of monolayer cells, but accumulates as a matrix between the cells and the dish. Laminin does, however, seem to share some properties with fibronectin. Both bind to proteoglycans⁸ and to collagen, but whereas fibronectin binds to all collagens, laminin only seems to interact with Type IV procollagen, the species of collagen which is found in basement membranes. Very little is known about how these properties of laminin influence its deposition *in vivo*. In the skin, the basal lamina is composed of two layers, the *lamina lucida* next to the epidermis and the

lamina densa next to the dermis. Immunofluorescence studies have localized laminin to the *l lucida*² and Type IV procollagen to the *l densa*⁹ suggesting that here, at least, there is no intimate intermingling of the two components. Alternatively, laminin may be present throughout the basal lamina but its antigenic sites masked in the *l densa*.

Although its almost ubiquitous distribution in basement membranes suggests that laminin plays an important role in epithelial tissue organization, is there any real evidence for this? Studies in which epithelial cells were given a choice of collagen matrix on which to attach showed that they bind slowly but preferentially to Type IV procollagen¹⁰. It now appears¹¹ that this binding specifically requires laminin. Normally the cells make it for themselves slowly, but attachment can be speeded up by adding purified laminin, an effect which is blocked by anti-laminin antibody. Laminin also appears to be the active component of various conditioned media which promote epithelial attachment and growth¹². Fibronectin cannot substitute for laminin in these assays, and, conversely, laminin does not mediate the binding of fibroblasts to collagen. These results suggest, but by no means prove, that at least one *in vivo* role of laminin is to help epithelial cells keep their feet firmly on the ground.

It is an important feature of epithelial sheets and tubes that the cells are asymmetrical, with distinct basal and apical surfaces. During embryonic development

there are numerous descriptions of epithelial tissues emerging from clusters of precursor cells which are not themselves polarized in this way, and it is of interest that in two cases the onset of laminin synthesis seems to correlate with the time of this polarization. One example is the morula stage mouse embryo³ which at the time of compaction undergoes transition into a group of cells with definite inner and outer surfaces. The second example is the aggregation of mouse embryo kidney mesenchyme into tubules¹³. In both cases laminin synthesis may only be an early manifestation of epithelial differentiation, but it is also possible that its assembly into a basement membrane actually directs cell polarization in some way.

Finally, what happens to basement membrane laminin when epithelial cells metastasize? In a recent paper, Hayman and colleagues¹⁴ showed that after virus transformation, rat kidney cells lose the ability to lay down laminin and fibronectin in an extracellular matrix, even though they continue to synthesize the two glycoproteins at the same or increased rates. The authors speculate that a matrix cannot be organized by the transformed cells because of a change in their surface glycosaminoglycans, to which, as we have seen, both laminin and fibronectin bind. However, it could also be argued that the primary lesion is in the polarized cytoskeletal framework of the cells which allows them to spread their feet out over a surface, and through transmembrane connections keeps the matrix firmly anchored down. □

- 1 Tampl *et al* *J Biol Chem* 254, 9933 (1979)
- 2 Foidart *et al* *Lab Invest* 42, 336 (1980)
- 3 Leivo *et al* *Dev Biol* 79, 100 (1980)
- 4 Howe & Salter *Dev Biol* 77, 480 (1980)
- 5 Hogan *et al* *Dev Biol* 80, 289 (1980)
- 6 Chung *et al* *Cell* 16, 277 (1979), Sakashita & Rouslahti *Archs Biochem Biophys* 205, 283 (1980)
- 7 Alitalo *et al* *Nature* 287, 465 (1980)
- 8 Sakashita *et al* *FEBS Lett* 116, 243 (1980)
- 9 Yaota *et al* *J Invest Dermatol* 70, 191 (1978)
- 10 Murray *et al* *J Cell Biol* 80, 197 (1979)
- 11 Terranova *et al* *Cell* 22, 719 (1980)
- 12 Vlodavsky & Gossardarowicz *Nature* 289, 304 (1981)
- 13 Ekblom *et al* *Proc natn Acad Sci U S A* 77, 485 (1980)
- 14 Hayman *et al* *J Cell Biol* 88, 352 (1981)



100 years ago

MR DARWIN ON VIVISECTION

"Dear Sir, — In answer to your courteous letter of April 7 I have no objection to express my opinion with respect to the right of experimenting on living animals. I have all my life been a strong advocate for humanity to animals, and have done what I could in my writings to enforce this duty. Several years ago, when the agitation against physiologists commenced in England, it was asserted that inhumanity was here practised and useless suffering caused to animals. It is right to add that the investigation of the matter by a Royal commission proved that the accusations made against our English physiologists were false. From all that I have heard however I fear that in some parts of Europe little regard is paid to the sufferings of animals, and if this be the case I should be glad to hear of legislation against inhumanity in any such country. On the other hand I know that physiology cannot

possibly progress except by means of experiments on living animals, and I feel the deepest conviction that he who retards the progress of physiology commits a crime against mankind. Any one who remembers, as I can, the state of this science half a century ago must admit that it has made immense progress, and it is now progressing at an ever-increasing rate.

"What improvements in medical practice may be directly attributed to physiological research is a question which can be properly discussed only by those physiologists and medical practitioners who have studied the history of their subjects, but, as far as I can learn, the benefits are already great. However this may be, no one, unless he is grossly ignorant of what science has done for mankind, can entertain any doubt of the incalculable benefits which will hereafter be derived from physiology, not only by man, but by the lower animals. Look, for instance, at Pasteur's results in modifying the germs of the most malignant diseases, from which, as it so happens, animals will in the first place receive more relief than man. Let it be remembered how many lives and what a fearful amount of suffering have been saved by the knowledge gained of parasitic worms through the experiments of Virchow and others on living animals. In the future every one will be astonished at the ingratitude shown, at least in England, to these benefactors of mankind. As for myself, permit me to assure you that I

honour, and shall always honour, every one who advances the noble science of physiology.

"Dear sir, yours faithfully,

"CHARLES DARWIN

Mr M G Mulhall sends us the following curious note, which we give without comment — "Although Shakespeare is supposed to have taken the idea of Hamlet from the Danish historian Saxo Grammaticus, there are such points of resemblance with the Arabic chronicle of Nigiaristan, respecting Montasser, tenth Caliph of Bagdad, that I venture to call your attention to the same. The points of analogy are as follows: 1 That Montasser is murdered by putting poison in his ear. 2 The ghost scene, in which his father appears to him. 3 The displaying of tapestry before the Caliph and his court, in which the tapestry represents a tragedy identical with the late Caliph's murder."

Mr H L Janssen van Raay writes to us from Batavia, March 21, that in the enumeration of the different geographical societies of the world in *Nature*, vol. xxiii p. 299, the Geographical Society at Samarang (Java), founded in 1879, was omitted.

From *Nature* 23, 21 & 28 April, 1880

ARTICLES

The Earth's radius and the G variation

V. M. Canuto

NASA Goddard Space Flight Center, Institute for Space Studies, New York, New York 10025 and
Department of Physics, The City College of the City University of New York, New York 10021, USA

It has long been assumed that if the gravitational constant G was larger in the past, the Earth's radius had to be smaller. The assertion holds provided the input from microphysics (in particular the equation of state) is independent of G . While this is true for some theories of gravity with variable G it is not so in the scale covariant theory, where the pressure can be affected by a variable G in a way that, for a constant mass of the Earth, a larger G in the past implies a larger Earth's radius. Comparison with recent palaeomagnetic data is presented.

MODERN observations¹ of increasing accuracy seem to confirm that the phenomenon of gravity is correctly understood in terms of Einstein's geometrical interpretation and satisfactorily described by the Einstein equations. There, therefore, seems to be no cogent reason to modify Einstein's theory as a description of gravity, as there is no reason to modify atomic theory. Within their limits, gravitational theory and atomic theory are complete and thus far satisfactory theories.

In cosmology, however, where gravitational phenomena involving large fractions of the age of the Universe are analysed not with gravitational but with atomic clocks, the two dynamics must be coupled. Let Δt_E (E for Einstein) be the time interval marked by a clock governed by gravitational dynamics (two orbiting planets, for example) and Δt the time marked by an atomic clock. Because the gravitational constant G and the total mass M may be thought of as the 'springs' of the gravitational clock, much in the same way as e , $\hbar$ and m may be considered the 'springs' of the atomic clock, then $\Delta t_E = f(G, M)$ and $\Delta t = g(e, \hbar, m)$.

To study the large times involved in cosmology a relationship between Δt_E and Δt as a function of time or equivalently of the age of the Universe is required. The lack of a unified theory of electromagnetic and gravitational forces, which would give us the answer, has been avoided so far by assuming that the ratio $\beta(t) = \Delta t_E / \Delta t$ at any time in the past was the same as today. This is known as the strong equivalence principle, SEP. While the adoption of the SEP has not yet led to flagrant disagreement with observations, it remains an unproved assumption, because its confirmation would require two measurements of $\beta(t)$ at two different times, the age of the Universe serving as a time scale. A more popular way of expressing a possible non-constancy of $\beta(t)$ is by saying that G may vary with respect to atomic time². In fact, a non-constant $\beta(t)$ is equivalent to one of the springs of the gravitational clock, say G , varying with respect to the atomic time.

Given the fundamental role of the SEP for the entire cosmological edifice, it is natural to expect that the nature of the function $\beta(t)$ be determined not by an *a priori* assumption, but by either a unified theory, which we do not have, or by observations of which we now have enough, covering a large fraction of the age of the Universe.

Cosmological data to put limits on β/β were used in ref 3 while the past luminosity of the Sun was analysed in refs 4 and 5. This article concentrates on the dependence of the radius of the Earth on $G = G(t)$.

The newtonian framework is shown to be unsuitable because it leads to inconsistencies (see refs 3, 4, 6, 7). Then it is shown that the Einstein framework leads, as expected, nowhere. Finally, the scale covariant theory is used through which the effects of a non-constant $\beta(t)$ can be treated and quantified consistently.

The Earth's radius

While the first analysis⁸ of the consequences of $G = G(t)$ dealt with the Sun's luminosity (the result $L_\odot \sim G^7$ has been shown⁴ to be $L_\odot \sim \text{constant}$, see also ref 5), more attention seems to have been given to the Earth's radius^{9,10}. The general consensus seems to be that if G was larger in the past, the Earth's radius had to be smaller (see, for example, refs 9–12).

To arrive at this qualitative conclusion, one uses (a) the hydrostatic equilibrium equation, (b) an equation of state and (c) the assumption of a constant total rest mass, ($\rho_0 = mn$, $n = N/V$, $M = mN$)

$$\frac{1}{\rho_0} \frac{dp}{dr} = -G \frac{m(r)}{r^2} \quad (1a)$$

$$p \propto \rho^\gamma \quad (1b)$$

$$M = \text{constant} \quad (1c)$$

thus deriving

$$R \propto M^{(\gamma-2)/(3\gamma-4)} G^{-1/(3\gamma-4)} \sim G^{-1/(3\gamma-4)} \quad (2)$$

which is the basic relation for all the qualitative conclusions that a larger G implies a smaller R , and therefore an expanding Earth.

While equations (1a), (1b) and (1c) form a set of consistent relations when G is constant, is the same true when G becomes a variable? The different exponentials of G and M in equation (2) are in sharp contrast with equation (1a) where G enters only in the combination GM . The asymmetry is introduced using an equation of state in the form (1b), which contains M but not G , the usual justification being that atomic relations are immune from whatever happens to G , which is valid only if gravity and atomic physics are kept separated. However, a variation of G represents a change of one dynamics with respect to the other, a process that can only be meaningful if we relate, not separate, the two dynamics. While the assumption that the two dynamics are disconnected may be logically consistent with the SEP, the purpose here is to relax the SEP allowing the two dynamics to depend on one another through the function $\beta(t)$ or equivalently $G(t)$. Furthermore, while we may reasonably expect an isolated hydrogen atom to be independent of β , this need no longer be true for a many-body quantity like an equation of state, which involves ingredients like the integration over phase space, the Liouville equation, as well as thermodynamic relations.

To show how equation (1b) may lead to incompatibilities with a $G = G(t)$, let us consider the equation of state $pV = NkT$, and use the adiabatic relation $TV^\Gamma = \text{constant}$, where $N = \text{constant}$, that is $M = \text{constant}$. We obtain $p \propto \rho^\gamma$ ($\gamma = 1 + \Gamma$), that is equation (1b) (equations (1b) and (1c) are, therefore, related). Furthermore, the basic ingredient in this derivation is the

adiabatic relation $TV^\Gamma = \text{constant}$, which comes from integrating the first law of thermodynamics, which in turn is a statement of conservation of energy

However, if $G = G(t)$, the potential energy $F = -G(t)m_2/r^2 = -\nabla V(r, t)$ becomes time dependent and we do not have energy conservation¹³. An inconsistency has surfaced.

The argument may be formalized by moving from the newtonian framework to the einsteinian one, not because general relativistic effects are relevant to the Earth, but because GR is the correct description of gravity. While the previous argument indicates that equations (1b) and (1c) are related, adopting the einsteinian framework will show that all the relations (1) are connected. In fact, the derivation of equation (1a) is based on the momentum conservation law, while equations (1b) and (1c) are related to the energy conservation law for the total energy density ρ . These laws can be combined into the conservation of the energy momentum tensor $T_{\mu\nu}$ (ref 14, equation 126.1), which in arbitrary coordinates reads

$$T^{\mu\nu}_{;\nu} = 0 \quad (3)$$

Let us now consider Einstein's equations,

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R, \quad G_{\mu\nu} = -8\pi GT_{\mu\nu}, \quad (4a)$$

$$G^{\mu\nu}_{;\nu} = 0 \quad (4b)$$

Equation (4b) is the Bianchi identity, a geometrical relation valid irrespective of the content of the right-hand side of equation (4a)¹⁵. Clearly equations (1), or the global form, equation (3), are compatible with equation (4b) only if G is constant.

This argument indicates that taking G in equation (1a) variable is inconsistent with the simultaneous use of equations (1b) and (1c) in which G is absent not because it has no *a priori* right to be there, but because it has already been assumed to be constant.

The Einstein framework

Conventional general relativity is based on the existence of a metric field which obeys the Einstein field equations (4b) with or without the right-hand side (ref 16 and P. Bergmann, personal communication), $GT_{\mu\nu}$, a quantity that Einstein¹⁷ considered "a formal condensation of all those things whose comprehension in the sense of a field theory is still problematic". Equations (4a) were constructed from left to right¹⁵. Equation (4b) means that the right-hand side of equations (4a) must have zero divergence. While one may choose G to be constant, thus yielding $T^{\mu\nu}_{;\nu} = 0$ which embodies the 'standard' equations of motion and the 'standard' energy conservation equation, this choice is clearly not dictated by the theory of gravity. For that reason Einstein considered his theory 'incomplete', in the sense that it did not prescribe how to construct the source term from within. Given equation (4b), it would have been equally possible to choose

$$(GT^{\mu\nu})_{;\nu} = 0 \quad (5)$$

with G as a variable. Equation (5) indicates that $T^{\mu\nu}_{;\nu} = 0$, equation (3), is no longer true, because there is now a source term proportional to G . Because the energy conservation equation has acquired new terms involving G , so will the resulting equation of state. To summarize, as the energy-momentum tensor $T_{\mu\nu}$ is a function of the pressure p and the total energy density ρ , we have from equation (4) the two possibilities

G constant plus Bianchi identity $\rightarrow T^{\mu\nu}_{;\nu} = 0 \rightarrow p = p(\rho_0)$, $M = \text{constant}$

G variable plus Bianchi identity $\rightarrow (GT^{\mu\nu})_{;\nu} = 0 \rightarrow p = p(\rho_0, G)$, $M \neq \text{constant}$

independently of the specific form of $T_{\mu\nu}(p, \rho)$. To illustrate the modification to the p versus ρ relation, consider the following form for $T_{\mu\nu}$

$$T_{\mu\nu}(p, \rho) = (p + \rho)u_\mu u_\nu - pg_{\mu\nu} \quad (6)$$

By introducing the variables $p_* = pG$ and $\rho_* = \rho G$, we formally

reduce equation (5) to the standard form $T^{\mu\nu}_{;\nu} = 0$. Because G has 'disappeared', we recover again, as in the standard case $p_* \sim (\rho_0)_*$ and $M_* = \text{constant}$. Returning to the variables p and ρ_0 , we obtain finally

$$p \propto \rho_*^\gamma G^{\gamma-1} \quad (7a)$$

$$MG = \text{constant} \quad (7b)$$

Alternatively, as $T^{\mu\nu}_{;\nu} = 0$ implies $p_* dV + dU_* = 0$, on integrating, $(pV = NkT = \Gamma U)$, $NTGV^\Gamma = \text{constant}$. However, the zero pressure limit of $T^{\mu\nu}_{;\nu} = 0$ implies $nVG = NG = \text{constant}$, so that substituting T in $pV = NkT$, we recover equation (7a) (see also ref 4).

As for the hydrostatic equilibrium relation, the introduction of the new variables p_* and ρ_* has no effect on its structure, because G depends only on time. Equation (1a) is therefore consistent with a varying G .

Inserting equation (7a) into equation (1a), we derive,

$$R \sim M^{(\gamma-2)/(3\gamma-4)} G^{-1/(3\gamma-4)} G^{(\gamma-1)/(3\gamma-4)} \\ \sim (GM)^{(\gamma-2)/(3\gamma-4)} \sim \text{constant} \quad (8)$$

a result radically different from equation (2).

Alternatively, we can say that because only the starred quantities are physically meaningful, equations (1) should read ($\rho_0 = \rho$ for ease of notation)

$$\frac{1}{\rho_*} \frac{dp_*}{dr} = -\frac{m_*}{r^2}, \quad p_* \propto \rho_*^\gamma, \quad M_* = \text{constant} \quad (9)$$

which clearly show that G does not appear.

Scale covariant framework

The previous analysis has indicated that while it is in principle possible formally to make room for a variable G within the Einstein framework, the results turn out to be identical to those with a constant G . The message is that the quantity G has no 'identity' in such framework, thus confirming that there is no useful way of talking about a variable G within the Einstein theory.

To incorporate fully the possibility that atomic and gravitational times may differ over large fractions of the age of the Universe a formalism, the scale covariant theory (SCT) of gravity, was devised. A gauge function $\beta(t)$, epitomizing our ignorance of how gravity and atomic forces couple, was introduced through the equation

$$\Delta t_E = \beta(t) \Delta t \quad (10)$$

While the strong equivalence principle assumes $\beta(t) = \text{constant}$, a broad-based analysis of observational data ranging from astrophysics to cosmology (up to $z \approx 10^3$), gave no indication that $\beta(t)$ has to be constant^{3,4,6,7,18}. While this does not mean that $\beta(t)$ must be variable, it is a significant result once we consider that it was firmly believed that even simple data like the luminosity of the Sun would have clearly opposed the variability of $\beta(t)$. Furthermore, recent measurements of the period of the Moon by both atomic and gravitational clocks do not seem to yield the same result, the residual corresponding to

$$\frac{\dot{\beta}}{\beta} = (2.75 \pm 0.64) \cdot 10^{-11} \text{ yr}^{-1} \quad (11)$$

Gravitational times are defined as those times with respect to which the Einstein equations retain exactly their form as do all the purely gravitational expressions ensuing therefrom. However, with respect to atomic units, the gravitational equations take the form (ref 6, equation (2.10))

$$G_{\mu\nu} + f_{\mu\nu}(\beta) = -8\pi GT_{\mu\nu}(\beta) \quad (12)$$

The final aim of the SCT is to construct a viable two-times theory, within which the commonly assumed identity between atomic and gravitational times may be broken, $\Delta t \neq \Delta t_E$. For that to occur, the lagrangian $\mathcal{L}$ describing gravity and matter must not be scale invariant. While it is not *a priori* clear which of the

terms in $\mathcal{L}$ should break the symmetry, the choice is not large. Brands and Dicke (BD) chose the gravitational part $\mathcal{L}_g$, leaving $\mathcal{L}_m$, the matter part, unaltered. This implies that their $T_{\mu\nu}$ does not depend on β : the conservation law (3) holds unchanged, as do relations (1). Relation (2) is therefore consistent with the BD framework. However, the change in $\mathcal{L}_g$ means that the Einstein equations are modified by the addition of terms which coincide with our function $f_{\mu\nu}(\beta)$ only if the BD parameter ω is $-3/2$. The potentially most serious source of troubles facing a theory that alters the Einstein equations is that the expressions for the three fundamental tests acquire extra contributions which may spoil the ever increasing agreement with observations. As this seems to be the case in the BD theory, the SCT was constructed with the opposite point of view. The scale breaking terms were relegated entirely to $\mathcal{L}_m$, while the gravitational part of $\mathcal{L}$ was constructed in a scale invariant manner. The result is equation (12), where the function $f_{\mu\nu}(\beta)$ assures that the left-hand side of equation (12) is scale invariant, that is it has power zero under the scale transformation $ds \rightarrow ds' = \beta(x) ds$, where for any quantity A , we have $A \rightarrow A' = \beta^{\pi(A)} A$, $\pi(A)$ being the 'power' of A .

This approach has been questioned¹⁹ on the grounds that the 'presently accepted microscopic lagrangian' can be shown to be scale invariant (a more detailed discussion is given elsewhere). One of the critical quantities entering the analysis of ref. 19 is $\pi(m)$, the power of a microscopic mass m . From the fact that $\hbar/mc$ is a length, it follows that $\pi(l) = \pi(\hbar) - \pi(m) = +1$. The claim¹⁹ that $\pi(m) = -1$ can only be made if $\pi(\hbar) = 0$, which in turn implies $\pi(e) = 0$, because $e^2/\hbar c$ is a pure number. Now because the electric charge enters only in the lagrangian term representing interactions, any assumption about $\pi(e)$ is an assumption about the scaling property of the interaction. Because of the general relation $\beta GM = \text{constant}$ (see equation (14)), $\pi(m) = -1$ implies $g = 2$. Now considering the electromagnetic tensor $F^{\mu\nu}$, if the contribution to $T_{\mu\nu}$ from the free electromagnetic field is taken to be scale invariant then the power of $F^{\mu\nu}$ is $-3 - g/2$ which is -4 . Because $J^\mu = enu^\mu$ has power -4 , the interaction equation $F^{\mu\nu}{}_{;\nu} = J^\mu$ is scale invariant. As such, it may not hold in the SCT because the final aim is that of allowing atomic clocks to scale differently from gravitational clocks. Because the gravitational part $\mathcal{L}_g$ has been constructed in a scale invariant form (hence scale invariant theory of gravitation) with the result that the period of a planet with respect to atomic clocks scales like $P \sim \beta^{-1}$ (ref. 6, equation (4.19)), clearly the dynamic equation governing an atom cannot be $F^{\mu\nu}{}_{;\nu} = J^\mu$ because being scale invariant, it would yield $P(\text{atom}) \sim \beta^{-1}$ too, thus leading to no difference between atomic and planetary dynamics, that is to a constant G . If we accept 'the presently accepted form of $\mathcal{L}_m$ '¹⁹ (an equation like $F^{\mu\nu}{}_{;\nu} = J^\mu$) then clearly the only possibility for a varying G is to go back to $\mathcal{L}_g$ and to devise a modified BD approach. This line of reasoning, which does not contemplate the possibility advocated by the proponents of the SCT, cannot be taken to imply nor to have proved that either (1) G is constant or (2) that for it to vary, the only avenue open is $\mathcal{L}_g$. Progress in modifying $F^{\mu\nu}{}_{;\nu} = J^\mu$ using gauge fields and neutral current concepts, has recently been made (P. J. Adams and J. Anderson, personal communication). Far from being mutually exclusive, the approach proposed in ref. 19 and the point of view of the SCT are in our opinion but different approaches towards the same goal.

Let us now prove that βGM is constant. It has been shown⁶ that the left-hand side of equation (12) has zero co-covariant derivative (indicated by * instead of ,). We therefore have, using equation (A.19) of ref. 6 and the fact that $G_* = 0$,

$$T^{\mu\nu}{}_{;\nu} + (\pi + 6)\phi_{,\nu} T^{\mu\nu} - \phi^{\mu} T^{\nu}{}_{;\nu} = 0 \quad (13)$$

with $\phi \equiv \ln \beta$, $\phi_{,\nu} \equiv \phi_{,\nu}$. The quantity π is the power of $T^{\mu\nu}$. Multiplying equation (13) by u_μ and using for $T_{\mu\nu} = \rho_0 u_\mu u_\nu$, we obtain on integrating ($\rho_0 V = M$)

$$\beta GM = \text{constant} \quad (14)$$

where $V/V = u^\mu{}_{;\mu}$, and $\pi = \pi_{S1} = -4 - g$, $g = \pi(G)$. We have

required that the right-hand side of equation (12) be scale invariant, $\pi(GT_{\mu\nu}) = 0$. Equation (14) makes it clear that the power of m must be chosen consistently with the dynamical constraints of the theory, and not assumed *a priori*. In fact because $\pi(m) = 1 - g$, a $G\beta^2 = 1$ gauge yields $\pi(m) = -1$, whereas $G\beta = 1$, the so-called non-matter creation case, yields $\pi(m) = 0$.

The very fact that in equation (12) $T_{\mu\nu}$ is taken to depend on β implies, even without further qualifications, that we cannot use, without checking their validity, any of the standard relations (1a), (1b) and (1c) because, strictly speaking, they are valid only for constant β . The β dependence of microscopic relations, marking a sharp contrast with the BD theory, is the price one must pay for constructing a theory with a built in two-times scheme and which at the same time preserves the agreement between observations and predictions for the fundamental tests, whose expressions in atomic units can be derived by direct scaling of the corresponding gravitational expressions. Because at any given time the scale function $\beta(t)$ can be normalized to unity, a set of measurements at one time cannot reveal the presence of β . In this sense, the SCT theory has preserved agreement with the three tests⁷. The scaling of gravitational expressions to obtain the corresponding atomic ones is all that is required when dealing with purely gravitational phenomena, like periods or distances of planets, which by assumption are given by Einstein equations and are constant with respect to gravitational clocks. On the other hand, non-gravitational quantities must be derived consistently with equation (13), so that the influence of the function β can be fully accounted for. This is a major endeavour because it calls for a reconstruction of microscopic physics, a process which, while accounting for the fact that β can enter at any stage, must satisfy the following two constraints: (1) the atomic clock ensuing from such theory must not depend on β , (2) the many-body behaviour of the theory must yield results consistent with the classical ones derived from equation (13). A typical example of such 'matching' relation is the continuity equation, $\rho + \rho \text{div } J = 0$, which can be obtained both from equation (3) and from the Schrodinger equation. These requirements clearly indicate that the attempt at constructing such theory is far more complex and difficult than that faced by Brans and Dicke. Descending from the Einstein equation to microphysics (thus running in the opposite direction to the traditional one, which assumes the validity of the microscopic scheme and makes connection with gravity by going from a flat to curved space) has to my knowledge never been worked out in detail. In spite of this, we have already derived⁶ the influence of β on several microscopic relations, and the corresponding observational implications such as the energy versus frequency for a free photon, the adiabatic scaling law for a perfect gas, the equilibrium radiation versus frequency and temperature expression, and the luminosity versus distance relation (see also ref. 5). No contradictions have yet appeared.

Having clarified this point, let us return to the original problem of evaluating R as $R(t)$. The full problem is very complex because we do not yet possess a full description of atomic interactions which determine the equation of state for an object like the Earth. We can therefore present only a hopefully good approximation to the exact solution. Returning to equation (13), and using for $T_{\mu\nu}$ equation (6), we first separate out the rest mass contribution from the total energy density ρ , $\rho = \rho_0 + u$. The rest mass part of equation (13) can be easily integrated. The result is equation (14). The $\rho - \rho_0 = u$ part satisfies the equation ($U = uV$)

$$dU + p dV + [(1 - g)U + 3pV] d\phi = 0 \quad (15)$$

which is the first law of thermodynamics generalized to include the β terms³. Let us now write

$$p = \frac{N}{V} kT \left[1 + B(T) \frac{N}{V} + C(T) \left(\frac{N}{V} \right)^2 + \dots \right] \quad (16)$$

where $B(T)$ and $C(T)$ are the so-called virial coefficients²⁰. We first consider the perfect gas case, $B = C = 0$. Using $pV = NkT$

and $U = NkT/\Gamma$ and remembering that $N \sim (\beta G)^{-1}$, we integrate equation (15), with the results ($\gamma = 1 + \Gamma$)

$$p \propto \rho_0^\gamma \left(\frac{G}{\beta^2} \right)^{\gamma-1} \quad (17a)$$

$$T\beta^{3\Gamma} \sim \frac{1}{V^\Gamma} \quad (17b)$$

For a more general derivation, valid for non-adiabatic transformations, see ref 5

Note that there is no choice of gauge, a G versus β relation, that reduces equations (17a) and (14) to equations (1b) and (1c), at least for $\gamma \neq 1$. In fact, to recover equation (1b), we must choose $G \sim \beta^2$, which implies $\beta < 0$, contrary to equation (11). This choice would also imply $M \sim \beta^{-3}$, not $M \sim \text{constant}$. To see the consequences of equation (17a), we use it in equation (1a), which has already been shown⁶ to be valid in the present framework

$$R \sim M^{(\gamma-2)/(3\gamma-4)} G^{-1/(3\gamma-4)} \left(\frac{G}{\beta^2} \right)^{(\gamma-1)/(3\gamma-4)} \sim \frac{1}{\beta} \quad (18)$$

on using equation (14). The result is independent of γ . We therefore conclude that

$$\frac{R}{R_0} = -\frac{\beta}{\beta_0} = -(2.75 \pm 0.64) 10^{-11} \text{ yr}^{-1} \quad (19)$$

where we have used equation (11). The predicted value of R , 400 Myr ago, is therefore ($R_0 = 1$)

$$R = 1.0110 \pm 0.002 \quad (20)$$

which can be compared with the recent palaeomagnetic result¹²

$$R = 1.020 \pm 0.028 \quad (21)$$

At first sight, the close agreement can be taken as justification for the approximations made to arrive at equation (18). However, a closer scrutiny reveals that the result must be taken with great care. First, as stressed in ref 12, an equation of state of the form of equation (17a), is not valid at the surface of the planet where $p = 0$ but $\rho_s \neq 0$. The criticism is clearly independent of G and β . More important, however, is that the result $R \sim \beta^{-1}$ implies that the radius of the Earth in einstein units $R_E = \beta R$ is constant. While possible, this result is not obvious because atomic quantities like e and m , which enter the relation p_E as $p_E(\rho_E)$ and therefore R_E , are not constant in einstein units.

Having discussed the limitations of equation (17a), we go back to equation (16) and include the interaction term $B(T)$ in the form $B(T) = -BT^{b-1}$. We simulate in this phenomenological way an attractive potential. Although equation (17b) is valid when $B = 0$, it is used because we are performing an approximation to the first term. Equation (16) then becomes ($\rho_0 = \rho$ for ease of notation)

$$p = A\rho^\gamma \left(\frac{G}{\beta^2} \right)^{\gamma-1} \left[1 - \frac{B}{A} \rho^{\tilde{\gamma}} \left(\frac{G}{\beta^2} \right)^{\tilde{\gamma}-1} \right] \quad (22)$$

where $\gamma = 1 + \Gamma$, $\tilde{\gamma} = \gamma_* - \gamma$, $\gamma_* = 2 + (\gamma - 1)b$. An equation of state for the Earth of the type (22) (for $G = 1$ and $\beta = 1$) has been proposed by Birch²¹ with $\gamma = 7/3$ and $\gamma_* = 5/3$. While equation (17a) allows an exact solution, equation (18), equation (22) does not. For our purposes

$$R \sim \frac{1}{\beta} \left[1 - \frac{B}{A} \bar{\rho}^{\tilde{\gamma}} \left(\frac{G}{\beta^2} \right)^{\tilde{\gamma}-1} \right]^{1/(3\gamma-4)} \quad (23)$$

where $\bar{\rho}$ is an average (but time dependent) density. We further derive

$$\frac{R}{R_0} = -r \frac{\beta}{\beta_0}, \quad r - 1 = \frac{(2+g)\alpha}{\mathcal{D} - 3\tilde{\gamma}\alpha} \quad (24)$$

where $\mathcal{D} = 3\gamma - 4$, $\alpha = x(1-x)^{-1}$, $x = (B/A)\bar{\rho}^{\tilde{\gamma}} (G/\beta^2)^{\tilde{\gamma}-1}$. For the values of γ and γ_* given by Birch, we have $r - 1 = (2+g)\alpha/(3+2\alpha)$, which for $g = 1$ ranges from zero ($\alpha = 0$) to $3/2$ ($\alpha =$

∞). The first value corresponds to equation (18), $\alpha = \infty$ implies $x = 1$, that is $\bar{\rho} = \rho_s$, not acceptable. If p must vanish at $\rho = \rho_s$, then $x = (\bar{\rho}/\rho_s)^{\tilde{\gamma}}$. In spite of the large excursion in α , the corresponding variation in $r - 1$ is only from 1 to 1.5. Taking $\alpha = 1$ ($x = 1/2$, corresponding to $\bar{\rho} = 8 \text{ g cm}^{-3}$, $\rho_s = 2.84 \text{ g cm}^{-3}$) and $g = 1$, we have $r = 1.6$, so that 400 Myr ago

$$R = 1.018 \pm 0.004 \quad (25)$$

in better agreement with equation (21). Finally equation (24) may alternatively be written, using equation (14),

$$R \sim (GM)^r, \quad \frac{R}{R_0} = r \frac{(GM)}{GM} \quad (26)$$

to be compared with equation (2). As r is positive (at least for $g + 2 > 0$) the SCT predicts a result qualitatively different from equation (2) for a constant mass M , a larger G implies a larger radius, not a smaller one.

The method used to arrive at equation (22) is not meant to be a microscopic derivation of the p versus (ρ, β) relation for the Earth, but only an heuristic approach to show how β may enter the problem. Even without β , it is difficult to derive²² an equation of state for the Earth. Because all the $p = p(\rho)$ relations discussed in ref 22 contain no β and because $\beta = 1$ corresponds to einstein units, one might be tempted to think that the relations in ref 22 are actually valid in gravitational units and that a simple scaling could then yield the corresponding atomic relations. This is, however, not the case. Such a procedure would yield a relation where the difference of the power of ρ and that of (G/β^2) is unity, like the first term of equation (22). In turn this yields $r = 1$. It can be seen that $r - 1$ is directly proportional to the difference of these two exponents. Furthermore, $r = 1$ implies $R_E = \text{constant}$ producing difficulties already discussed. With the present procedure, this does not happen. The powers of ρ and that of (G/β^2) of the second term of equation (22) do not differ by one but by two. An extra term $(G/\beta^2)^{-1}$ appears which makes $r \neq 1$. This extra factor, which breaks the symmetry that one would invariably get out of a simple scaling process, cannot be arrived at from gravitational units where $\beta = \text{constant}$.

Lunar data

Having derived equation (24) but having no estimates for β/β_0 from within the theory, we can either use palaeomagnetic data for R/R_0 and thus determine β/β_0 (and possibly G/G_0), or we evaluate β/β_0 from other data and then compare the resulting R/R_0 with palaeomagnetic data. The only other data available are the rate of change of the Moon's period measured in both atomic and gravitational units: these are ($P = 2\pi/n$ is the period of revolution of the Moon)

n_E (gravitational time) (arc s cyr ⁻²)		n (atomic time) (arc s cyr ⁻²)	
-26.0 ± 2.0	ref 23	-21.4 ± 2.6	ref 26
-28.5 ± 3.1	ref 23	-23.6 ± 1.5	refs 26, 27
-30.0 ± 3.0	ref 7	-24.6 ± 3.9	refs 26, 28
-27.4 ± 3.0	ref 24		
-30.6 ± 3.1	ref 25		

where, because of equation (10), $n = \beta n_E$. A least-square fit analysis yields (β/β_0 in 10^{-11} yr^{-1}), note that $1 \text{ cyr} = 10^3 \text{ yr}$

$$\text{All equations} \quad n_E = -27.97 \pm 0.78, \quad \beta/\beta_0 = 2.75 \pm 0.64 \quad (28a)$$

$$\text{First } n_E \text{ deleted} \quad n_E = -29.11 \pm 0.73, \quad \beta/\beta_0 = 3.41 \pm 0.54 \quad (28b)$$

$$\text{Second } n_E \text{ deleted} \quad n_E = -27.87 \pm 0.91, \quad \beta/\beta_0 = 2.69 \pm 0.72 \quad (28c)$$

Because the value of β/β_0 retaining all the eight equations turns out to fall between the other two, we have adopted this value, equation (11). Finally, independently of the way we determine β/β_0 (palaeomagnetic or lunar data), the value of G/G_0 is gauge dependent and is to be determined using the relations $G/G_0 = -g\beta/\beta_0$, $\pi(m) = 1 - g$

Moment of inertia

Using equation (24), and $M = \text{constant}$ we obtain $I/I = -2r\beta/\beta$, or equivalently $I/I = 2rG/G$, with $r > 0$. This result agrees with equation (A 16) of ref 29, because the G in equation (A 1) must be written as $G_* = G(\beta^2/G)^{\gamma-1} = G^{4-3\gamma}$, due to equation (17a). If so, equation (A 16) becomes $I/I = aG/G$, with $a = (3\gamma - 4)/10 > 0$.

This result is, however, of limited use because it is in atomic units and it represents only the cosmological and not the total contribution to I/I , whereas the estimates reported in the literature are total variations—due to all possible causes, cosmological and otherwise—and they are all with respect to gravitational units. To prove this last point, we must remember that the basic ingredient (see refs 30, 31) is the conservation of the total angular momentum for the Earth–Moon system $J_E + L_E = \text{constant}$. If the Earth is considered a point mass, this implies the conservation of the orbital angular momentum L_E , a law valid only with respect to gravitational units^{6,7}. In atomic units L scales like $L \sim \beta^2/G$ ($\sim G^{-3}$ for constant mass) (ref 6, equation 4.7).

We now consider I_E/I_E from all sources and expressed in gravitational units. While growth lines in fossils promised to be a valuable source of information³⁰, recent difficulties in interpretation have indicated that quantitative results are not yet possible³². We shall therefore rely on ancient astronomical data (see Chapter 10, ref 31), and take equation (10.2.2b) from ref 31 and generalize it to include a non-constant I_E . We obtain

$$\frac{I_E}{I_E} = -\frac{\Omega_E}{\Omega_E} + \frac{\gamma}{3} \frac{n_E}{n_E} + H \quad (29)$$

where $\gamma = (1+b)L_0/J_0$, b accounting for the solar contribution. H represents the contribution of the time variation of dynamical elements describing the size, shape and orientation of the Earth–Moon system. If $J = I\Omega \cos \phi$, and $L = M_e M_m / (M_e + M_m) \times R^2 n (1-e^2)^{1/2} \cos i$, then $H = H(e, \phi, di/dt)$ (see ref 33).

Ancient astronomical observations can be used³¹ to construct a relation of the form $X_i n_E + Y_i \Omega_E = f$, where X_i and Y_i are numerical coefficients and f is a function of the observed less the computed position of the recorded event. As stated by Lambeck³¹ “the computed positions and locations” are arrived at “assuming purely gravitational motion”, further substantiating our claim that the ingredients are the standard newtonian equations which hold only with respect to dynamical units with G_E and M_E constant. Lambeck summarizes the results as follows

$$n_E = -28 \pm 2, \quad \Omega_E = -1,100 \pm 100 \quad (30)$$

This value for n_E is not substantially different from the one determined in equation (28). Using $L_0/J_0 = 4.89$ and $b = 0.21$, we obtain from equation (29)

$$\frac{I_E}{I_E} = -(0.086 \pm 0.031) 10^{-9} \text{ yr}^{-1} + H \quad (31)$$

To complete the computation we must now evaluate and then subtract from equation (31) the purely cosmological contribution. This is done by remembering that $I_E \sim M_E R_E^2 \sim \beta^2 R^2$, and using equation (24)

$$\left(\frac{I_E}{I_E}\right)_{\text{cosm}} = 2 \frac{(\beta R)}{\beta R} = -2(r-1) \frac{\beta}{\beta} = -(0.033 \pm 0.008) 10^{-9} \text{ yr}^{-1} \quad (32)$$

The difference

$$\frac{I_E}{I_E} - \left(\frac{I_E}{I_E}\right)_{\text{cosm}} = -(0.053 \pm 0.032) 10^{-9} \text{ yr}^{-1} + H \quad (33)$$

may now be attributed to geophysical phenomena such as the formation of the Earth's core³⁰. For the evaluation of H see ref 33.

Conclusions

The question whether a viable theory of gravitation exists that permits equations (1a), (1b) and (1c) as they are, while G is taken to vary, has been investigated with the following results.

Standard Einstein theory can be formally generalized to include a non-constant G . While equation (1a) remains valid, both equations (1b) and (1c) change in such a way that the final result is $R = \text{constant}$, demonstrating that G is actually a non-entity within the strict framework of Einstein.

In the BD theory, as by construction the matter lagrangian is not changed, equations (1) and (2) both hold. However, it is believed that the BD theory faces difficulties, at least in its present form.

Finally, while equation (1a) holds unchanged in the SCT equations (1b) and (1c) cannot be simultaneously satisfied. The most interesting result, however, which transcends the particular example equation (22), is that microphysics may be affected by a variable G . This may seem unusual at first as microphysics contains no G thus creating the impression that it ought to be immune from whatever happens to G . However, relations like the Boltzman equation, the radiative transfer equation, and the Navier–Stokes equations are derivable from and are consistent with the conservation law $T^{\mu\nu}_{;\mu} = 0$ for the energy momentum tensor, a law which holds only if G is constant, as equation (13) shows. The absence of G is actually a consequence of having chosen it to be constant.

More general derivations of classical non-gravitational relations (such as, Liouville equation⁵, Boltzman distribution⁵, radiative transfer equation⁵ and classical thermodynamics^{3,6}) in which β does explicitly appear, are possible, if constructed consistently with the general conservation laws that follow from the Einstein equation. For example, an equation of state of the form (17a) with $\gamma = 4/3$ holds^{3,6} for radiation (ρ_0 is replaced by the photon number density n_γ). Because $3p_\gamma = \rho_\gamma$ and $n_\gamma T \sim \rho_\gamma$, $\rho_\gamma \sim (\beta^2 G^{-1}) T^4$. It then follows that the Sun's luminosity $L_\odot$ scales like $L_\odot \sim \beta^{-1}/k$, where k is the opacity. For constant k and constant rest mass, $L_\odot \sim G$, contrary to the early statements that $L_\odot \sim G^7$ (for details, see refs 4, 5).

Our comparisons of the theoretical predictions with observations must be treated cautiously. The value of β/β derived from lunar data and used to estimate equations (20) and (25) is still subject to uncertainties due to the difficulties in estimating the errors, there are also doubts concerning palaeomagnetic data, or more precisely the assumptions made to arrive at equation (21) (ref 10, see, however, ref 34). Finally the evaluation of the function H , equation (33), is still fraught with large errors (see ref 33).

We prefer, therefore, to focus the dependence of R on G . If we write, in atomic units,

$$\frac{R}{R} = \left(\frac{R}{R}\right)_{\text{cosm}} + \left(\frac{R}{R}\right)_{\text{othersources}} \quad (34)$$

the present work predicts a reversal (compared with the standard treatment, equation (2)), of the sign of the first term in equation (34) when written in terms of G/G , equation (26).

This is in turn due to the new effect represented by the presence of G in the $p = p(\rho)$ relation. The G in the right-hand side of equation (1a) is overcompensated by the $G^{\gamma-1}$ factor in equation (17a), thus causing a larger G to correspond to a larger radius. However, the evaluation of the second term in equation (34) is outside the scope of this article, and we cannot conclude from our work alone that R is actually decreasing.

I thank Drs G. Blake, D. J. Stevenson and J. Anderson for constructive criticism, also Drs T. C. Van Flandern, S.-H. Hsieh, P. J. Adams, I. Goldman, J. R. Owen, K. M. Creer and J. D. Mulholland for assistance during the preparation of this manuscript.

Received 8 July 1980, accepted 9 February 1981

- 1 Will, C M in *General Relativity* (eds Hawking, S W & Israel, W) (Cambridge University Press, New York, 1979)
- 2 Dirac, P A M *Nature* **139**, 323 (1937)
- 3 Canuto, V M, Adams, P J, Hsieh, S-H & Tsang, E *Astrophys J Suppl* **41**, 243 (1979)
- 4 Canuto, V M & Hsieh, S-H *Astrophys J* **237**, 613 (1980)
- 5 Canuto, M & Hsieh, S-H *Astrophys J* (in the press)
- 6 Canuto, V M, Adams, P J, Hsieh, S-H & Tsang, E *Phys Rev D* **16**, 1643 (1977) (Paper I), **41**, 263 (1979) (Paper III), **41**, 301 (1979)
- 7 Canuto, V M, Hsieh, S-H & Owen, R J *Mon Not R astr Soc* **188**, 829 (1979)
- 8 Teller, E *Phys Rev* **73**, 801 (1948)
- 9 Jordan, P *The Expanding Earth* (Pergamon, New York, 1971)
- 10 Carey, S W *The Expanding Earth* (Elsevier, Amsterdam, 1976)
- 11 Wesson, P S *Cosmology and Geophysics* (Adam Hilger, Bristol, 1978)
- 12 McElhinny, M W, Taylor, S R & Stevenson, D J *Nature* **271**, 316 (1978)
- 13 Bishop, W T & Landsberg, P T *Nature* **264**, 346 (1976)
- 14 Landau, L D & Lifshitz, E M *Fluid Mechanics* (Addison-Wesley, Reading, 1959)
- 15 Dirac, P A M *General Theory of Relativity*, 25-26, 43-45 (Wiley, New York, 1975)
- 16 Bergmann, P *Introduction to the Theory of Relativity* (Prentice Hall, Englewood Cliffs, 1960)
- 17 Einstein, A in *Autobiographical Notes* (ed Schlipp, P A) 71, (Open Court, Chicago, 1979)
- 18 Canuto, V M & Hsieh, S-H *Phys Rev Lett* **44**, 695 (1980)
- 19 Bekenstein, J D & Meisels, A *Phys Rev D* **22**, 1313 (1980)
- 20 Landau, L D & Lifshitz, E M *Statistical Physics*, 224 (Pergamon, New York, 1958)
- 21 Birch, F *J geophys Res* **57**, 227 (1952)
- 22 Stacey, F D, Brennan, B J & Irvine, R D *Geophys Surv* (in the press)
- 23 Morrison, L V & Ward, C G *Mon Not R astr Soc* **173**, 183 (1975)
- 24 Goad, C C & Douglas, B C *J geophys Res* **83**, 2306 (1978)
- 25 Lambeck, K *Phil Trans R Soc* **287**, 545 (1977)
- 26 Van Flandern, T C *Astrophys J* (in the press)
- 27 Dickey, J O, Williams, J G & Yoder, C F *EOS* **61**, 46, 939 (1980)
- 28 Calame, D & Mulholland, J D *Science* **199**, 977 (1978)
- 29 Nordtvedt, K & Will, C M *Astrophys J* **177**, 755 (1972)
- 30 Runcorn, S K *Nature* **193**, 311 (1962), **195**, 1248 (1962), **204**, 823 (1964)
- 31 Lambeck, K *The Earth's Variable Rotation Geophysical Causes and Consequences* (Cambridge University Press, 1980)
- 32 Scrutton, C T in *Tidal Friction and the Earth's Rotation* (ed Broshe, P) (Springer, New York, 1978)
- 33 Hipkin, R G in *Growth Rhythms and the History of the Earth's Rotation* (eds Rosenberg, G D & Runcorn, S K) (Wiley, New York, 1975)
- 34 Schmidt, P W & Clark, D A *Geophys J R astr Soc* **61**, 95 (1980)

Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA

David B. McKay & Thomas A. Steitz

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511, USA

The 2.9 Å resolution crystal structure of Escherichia coli catabolite gene activator protein (CAP) complexed with cyclic AMP reveals two distinct structural domains separated by a cleft. The smaller carboxy-terminal domain is presumed to bind DNA while the amino-terminal domain is seen to bind cyclic AMP. Model building studies suggest that CAP binds to left-handed B-type DNA, contacting its major groove via two α-helices. It is possible that the CAP conversion of right- to left-handed DNA in a closed supercoil, is what activates transcription by RNA polymerase.

REPRESSOR and activator proteins that regulate gene expression at the level of transcription recognize specific nucleotide sequences in double-stranded DNA. Although it has been suggested that specific recognition involves α-helices fitting into the major groove of B-DNA¹⁻³ or anti-parallel β-strands in the minor groove⁴, no experimental evidence exists concerning the actual mode of sequence-specific interaction between protein and double-stranded DNA. Also, the molecular mechanism by which gene activators promote the activity of RNA polymerase, thereby switching on genes, remains unknown. We present here the structure determination of a protein that binds in a sequence-specific manner to double-stranded DNA and propose a mechanism by which the catabolite gene activator protein is able to switch on the catabolite-sensitive genes.

The catabolite gene activator protein (CAP), also called the cyclic AMP receptor protein, functions in *Escherichia coli* in the regulation of several catabolite-sensitive gene operons^{5,6}. Regulation by CAP is exerted at the transcriptional level with cyclic AMP acting as an allosteric effector. In the presence of a sufficient concentration of intracellular cyclic AMP, cyclic AMP forms a complex with CAP which binds to specific DNA sites near the promoters of several operons and alters the rate of their transcription by RNA polymerase. In the lactose (*lac*)⁷ and arabinose⁸ operons, the cyclic AMP-CAP complex is a positive regulator—it potentiates transcription. In the galactose operon, the presence of two overlapping promoters for RNA polymerase makes the situation more complex, cyclic AMP-CAP is required for initiation of transcription from one promoter, but inhibits transcription from the other. Thus, in this case it apparently acts as both a positive and negative regulator of operon expression⁹.

The active form of CAP is a dimer of identical subunits of molecular weight 22,500 and 201 amino acid residues^{10,11}. Proteolytic cleavage studies are consistent with the CAP subunit

having two separate structural domains¹². Results from a variety of techniques¹²⁻¹⁷ suggest that CAP undergoes a significant conformational change on binding cyclic AMP.

To develop a structural basis for understanding (1) the cyclic AMP-induced allosteric transition in CAP, (2) the site-specific recognition of DNA by CAP, and (3) the mechanism of transcription activation by CAP, we have initiated crystallographic studies of CAP and its complexes with ligands. In this article we report the structure of the cyclic AMP-CAP complex at 2.9 Å resolution.

Structure determination

CAP protein was purified and crystallized in the presence of 0.5 mM cyclic AMP as previously described¹⁸. The crystals are orthorhombic, space group P2₁2₁2₁, *a* = 46.5 Å, *b* = 97.1 Å, *c* = 105.4 Å, with one dimeric CAP molecule per asymmetric unit.

Intensities of crystallographic reflections were measured by diffractometer, using the Wyckoff scan algorithm¹⁹. Friedel pairs of reflections were measured on native crystals and five heavy-atom derivatives, for derivative data sets only ~75% of the total reflections, consisting of those which were most intense in the native data, were measured. Heavy-atom positions were located and refined by conventional methods²⁰ and the correct enantiomorph of the heavy-atom positions was determined using anomalous difference Fourier²¹. A summary of heavy-atom derivative preparation and refinement statistics is presented in Table 1. Combined multiple isomorphous replacement and anomalous scattering phases were computed to 2.9 Å resolution and had an average figure of merit of 0.74.

Part of the electron density map is displayed in Fig. 1. Most of the polypeptide backbone is well ordered and can be traced unambiguously in both subunits, however, regions of local disorder are encountered at the amino termini, the carboxy

Table 1 The heavy-atom derivative preparation and refinement statistics

	No of heavy-atom sites	Minimum resolution (Å)	R_{sym}	Resolution limits (Å)									
				40 00	23 20	11 60	7 73	5 80	4 64	3 87	3 31	2 90	
Native		2.9	0.048										
MeHg(I)	5	2.9	0.063	$\frac{f_{\text{rms}}}{E_{\text{rms}} R_c}$	2.92	3.80	3.61	3.68	2.72	2.62	2.57	3.08	
				R_c	0.34	0.36	0.37	0.35	0.41	0.53	0.53	0.46	
MeHg(II)	3	2.9	0.044	$\frac{f_{\text{rms}}}{E_{\text{rms}} R_c}$	1.25	2.58	2.65	2.92	2.86	2.16	2.18	2.09	
				R_c	0.54	0.40	0.44	0.36	0.40	0.62	0.64	0.60	
K ₂ ReCl ₆ (I)	2	3.5	0.031	$\frac{f_{\text{rms}}}{E_{\text{rms}} R_c}$	1.75	1.22	1.29	1.70	1.15	0.82	0.84		
				R_c	0.37	0.59	0.66	0.63	0.71	0.79	0.98		
K ₂ ReCl ₆ (II)	3	3.5	0.032	$\frac{f_{\text{rms}}}{E_{\text{rms}} R_c}$	1.31	1.21	1.21	1.38	1.22	1.02	0.97		
				R_c	0.76	0.57	0.69	0.64	0.72	0.90	0.83		
DMM	5	5.0	0.041	$\frac{f_{\text{rms}}}{E_{\text{rms}} R_c}$	1.33	1.78	1.82	2.28	2.22				
				R_c	0.57	0.59	0.45	0.50	0.48				
			$\langle m \rangle$		0.95	0.93	0.90	0.91	0.83	0.75	0.70	0.65	Overall
			No of reflections		17	150	386	634	999	1383	2011	2578	8158

Derivative soaking conditions MeHg(I) 1 mM methylmercuri- β -mercaptoethanol, 50 mM potassium phosphate, 0.5 mM cyclic AMP, pH 8.0, 1 day MeHg(II) 0.1 mM methylmercuri- β -mercaptoethanol, 50 mM potassium phosphate, 0.5 mM cyclic AMP, pH 8.0, 1 day K₂ReCl₆(I) 1 mM K₂ReCl₆, 50 mM potassium phosphate, 0.5 mM cyclic AMP, 0.1 mM dithiothreitol, pH 8.0, 1–2 days K₂ReCl₆(II) 2 mM K₂ReCl₆, 50 mM 2[N-morpholino]ethane sulphonic acid, 50 mM KCl, 0.5 mM cyclic AMP, 0.1 mM dithiothreitol, pH 6.4, 2 days DMM a drop of dimethylmercury was placed in capillary with mounted crystal and allowed to equilibrate for 3 days $\langle m \rangle$ = mean figure of merit

$$R_{\text{sym}} = \frac{\sum |I_i - \bar{I}|}{\sum \bar{I}} \quad f_{\text{rms}} = \left\{ \frac{\sum f_H^2}{n} \right\}^{1/2} \quad E_{\text{rms}} = \left\{ \frac{\sum (F_{\text{PH}} - |F_P + f_H|)^2}{n} \right\}^{1/2} \quad R_c = \frac{\sum |F_{\text{PH}} - F_P| - f_H}{\sum |F_{\text{PH}} - F_P|}$$

where I = diffraction intensity, f_H = calculated heavy-atom structure factor amplitude, F_P = structure factor amplitude of native crystals, and F_{PH} = structure factor amplitude of derivative crystals. Summations are overall reflections, n

termini and a coiled structure in the smaller domain marked A and A'.

Approximate α -carbon positions were obtained by placing markers on map sections and measuring their coordinates. Polarity of the polypeptide chain was established from the tilt of the larger side chains on the α -helices. Although many side chains are clearly visible in this map, no attempt has yet been made to place the several fragments of amino acid sequence obtained by Schlesinger. Cyclic AMP was identified because it was not continuously connected with the polypeptide backbone, included the most dense feature in the map (presumably the phosphate) and was too large to be an amino acid side chain.

Subunit morphology

The CAP subunit is folded into two distinct structural domains (Fig. 2). The larger amino-terminal domain has approximate overall dimensions of $25 \times 30 \times 35$ Å. It contains ~135 amino acid residues and consists primarily of a pair of helices (A and B, Fig. 2) which lead into an antiparallel β -strand folded into an eight-strand β -roll, also, a 40 Å-long α -helix C connects the two domains and partially closes off one end of the interior pocket formed by the β -roll. This helix extends from the small domain to the distal extremity of the large domain rather than forming a short, proximal bridge between the two domains. The topology of the eight-strand β -roll, sometimes referred to as the 'Swiss roll', has recently been observed, with minor variations, in three protein structures: the 'shell' domain of tomato bushy stunt virus protein²², Southern bean mosaic virus protein²³ and influenza haemagglutinin⁴⁶.

The carboxy-terminal domain is smaller, with approximate dimensions $20 \times 20 \times 30$ Å. It has about 65 amino acid residues and contains three distinct α -helices, D, E and F, the D helix is ~24 Å long, the F helix ~22 Å long, whereas the E helix is

rather short. Much of the small domain consists of irregular structure, some of which is partially disordered. Although the electron density marked AA' in Fig. 1 and corresponding to the region circled in Fig. 2 is well above the background level of the solvent region, it is not sufficiently well defined to place α -carbon positions with complete assurance. Thus, the polypeptide path shown in Fig. 2 at the carboxy terminus of the D and F helices illustrates only the general folding of the subunit in these regions and the connections within the circled region may have to be altered when amino acid sequence data are fitted to the electron density map. As the number of amino acid residues in our model is very nearly that expected from amino analysis^{10,11}, no significant portion of the molecule can be missing from the model.

Although there is a distinct cleft between the two domains of the CAP subunit, noncovalent contacts occur between them that presumably stabilize the relative orientation of the domains. Both subunits contain one similar region of noncovalent contact between their large and small domains: the end loop connecting strands 4 and 5 of the β -roll in the large domain is in contact with a loop of coil in the small domain. In addition, in one of the subunits, there is a second point at which the carboxyl end of the D helix may be within noncovalent contact distance of strand 5 of the β -roll of the large domain.

Dimer morphology

As the structures of the two subunits of the CAP dimer differ, there is no unique dyad axis relating them (Figs 3–5). Qualitatively, the cleft of one subunit appears 'closed' to the point of extensive van der Waals' contact between the large and small domains while the cleft of the second subunit is 'open', that is, the two domains have different relative orientations in the two subunits. Quantitatively, if the two large domains are super-



Fig. 1 Part of the electron density map, 81 Å horizontal by 71 Å vertical by 10 Å depth. Only positive contours are shown, the contour interval is approximately one-sixth the maximum peak height. The dashed line circumscribes one CAP molecule. The two oblong lobes near the bottom of the map are the small domains of the molecule, the two helices running approximately vertically near the centre of the map are the C helices. A, A', coiled structure in the smaller domain, B, B', positions of tight intermolecular contact between small domains and a large domain of a neighbouring molecule.

imposed by least-squares minimization of differences in corresponding α -carbon positions, a subsequent rotation of the small domain of one subunit by 28° is required to bring it into coincidence with the small domain of the second subunit. The 'hinge' region of the molecule appears to consist of approximately four to five amino acids in the region connecting the C helix to the small domain. This kind of conformational difference between chemically identical subunits has been widely observed, such as in yeast hexokinase²⁴, tomato bushy stunt virus coat protein²², immunoglobulins²⁵ and alcohol dehydrogenase²⁶.

Interactions between the two subunits are through direct contact along virtually the entire length of the two C helices, and contact of the β -roll of one subunit with the central region of the C helix of the other. No direct contact between the two small domains is apparent.

Although the entire dimer does not possess an overall 2-fold axis, the two pairs of domains are each related by separate rotation axes. These two rotation axes have been obtained by computing the rotation required for least-squares minimization of the difference in corresponding α -carbon positions of each domain. The results are summarized in Fig. 3. The large domains appear to be related by 2-fold symmetry (computed rotation $179.2 \pm 1.1^\circ$), the small domains may have a small distortion from dyad symmetry (computed rotation $186.7 \pm 3.7^\circ$), the angle between the rotation axis relating the large domains and that relating the small domains is 13° . The extensive contact between large domains suggests that these two domains maintain the same relative orientation in the presence and absence of various ligands. The observed asymmetry is, therefore, probably generated by an alteration in the orientation of a small domain relative to the large domains.

Cyclic AMP binding

The electron density map contains density for two cyclic AMP molecules of equal and full occupancy per CAP dimer, as anticipated from data on cyclic AMP binding to CAP in solution²⁷. The cyclic AMP is completely buried within the interior of the β -roll of the large domain. The peak of highest electron

density in the cyclic AMP molecule and the entire map is presumed to be the phosphate. Further indication of the orientation of the nucleotide is provided by crystals soaked in a solution of 8-bromoadenosine-3', 5'-cyclic monophosphate (8-Br-cyclic AMP). The cyclic AMP seems to be oriented with the phosphate bound in the loop connecting strand 6 to strand 7 in the β -roll, with the nucleotide base in proximity of the two C helices. Whether the cyclic AMP makes direct contact with one or both of the C helices, and whether the two cyclic AMP molecules per dimer are bound identically cannot be determined conclusively at this stage. The finding²⁷ that cyclic AMP shows negative cooperativity in its binding to CAP in solution may explain the formation of the asymmetric dimer observed in the crystal.

CAP domain structure and function

The crystal structure of CAP shows that the subunits consist of two distinct structural domains which may be generally true for allosterically controlled gene regulatory proteins. Limited proteolysis studies of *E. coli* lac repressor²⁸⁻³⁰, bacteriophage λ repressor³¹ and CAP^{12, 13} suggest that these proteins have a small DNA-binding domain and a larger domain that binds allosteric effectors and/or holds the subunits together. In all three cases it now seems that there is little, if any, direct interaction between the two DNA-binding domains.

Proteolytic cleavage of the lac repressor in appropriate conditions yields two separate protein fragments^{29, 30}: a tetrameric core that binds inducers, and a smaller monomeric domain having 51-59 amino acids, depending on the protease used. The small domain retains nonspecific³⁰ and operator-specific³² DNA-binding activity but does not bind inducers. Similarly, proteolytic cleavage of the λ repressor yields a dimeric core devoid of DNA-binding activity and monomeric domains which bind both nonspecific and λ operator DNA³¹.

It has been demonstrated that limited proteolysis of CAP yields an amino-terminal fragment of subunit MW 12,500, which still forms dimers and binds cyclic AMP but has lost its cyclic AMP-dependent DNA-binding activity¹². Isolation of a second intact carboxy-terminal fragment and determination of its cyclic AMP- and DNA-binding activities have not been

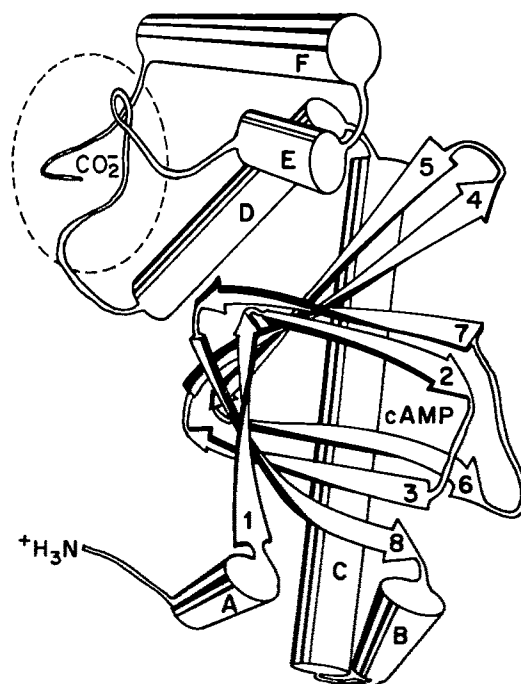


Fig. 2 Schematic drawing of the CAP promoter in the closed conformation. The numbered arrows represent β -sheet, the lettered cylinders α -helices. The binding site for cyclic AMP in the β -roll is labelled. The dashed line circumscribes the region which is not well ordered in the electron density map.

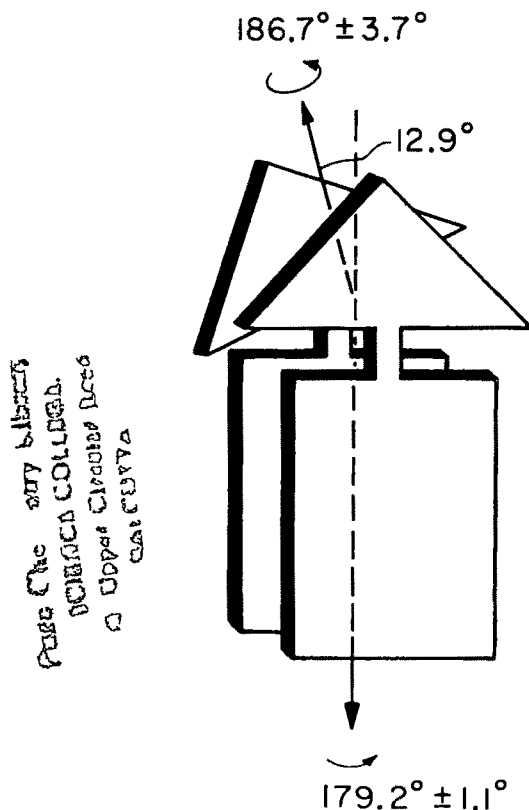


Fig. 3 Schematic diagram of the molecular symmetry of the CAP dimer. The rectangles represent the large N-terminal domains, the triangles the small C-terminal domains.

reported. Nevertheless, by analogy with the *lac* and λ repressors one might presume that the small domain of CAP binds to DNA while the larger domain binds the cyclic AMP and forms dimers.

Interaction of CAP with DNA

It has not proved possible to observe DNA binding to these crystals. Crystals soaked in solutions containing short (6–10 base pairs) oligodeoxyribonucleotides become substantially disordered. Because the small domains form tight intermolecular contacts with a large domain of a neighbouring molecule (regions B and B', Fig. 1), the most plausible DNA-binding surface of the small domains is inaccessible to oligonucleotides. Thus, direct study of DNA binding in this crystal form is probably impossible. To determine experimentally the structure of CAP bound to DNA, it will be necessary to co-crystallize CAP with an appropriate oligo-DNA.

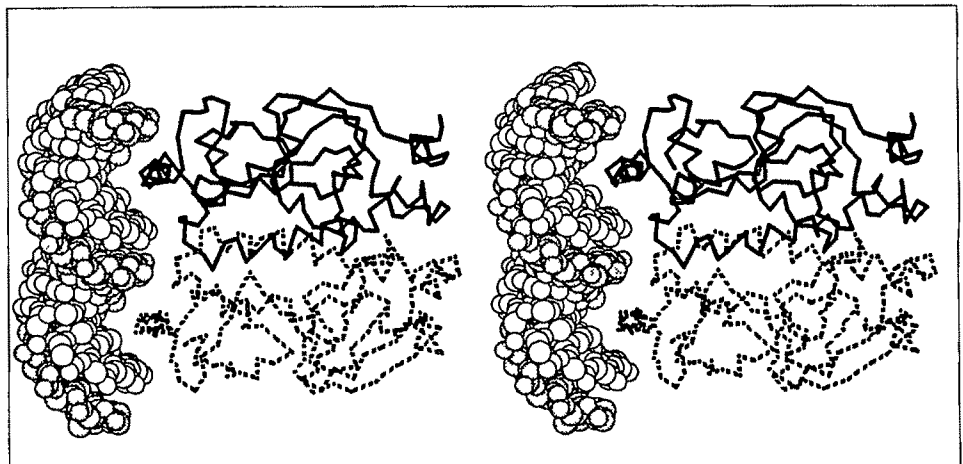
We have turned to model building to identify the DNA-binding site on CAP. Although we have not yet incorporated side chains into our CAP model, three general characteristics should be exhibited by the DNA binding site: (1) We expect a large degree of general structural complementarity between DNA and the potential binding site; (2) The correct CAP–DNA model must provide a structural basis for explaining the chemical protection^{33,34} and 'footprinting' (A. Schmitz, personal communication) data obtained on the CAP site of the *lac* operon; (3) Because the sequence of the CAP binding site at the *lac* promoter has a high degree of 2-fold symmetry, CAP probably binds to DNA with the approximate 2-fold axis that relates the two small domains coincident with the approximate 2-fold axis in the CAP site. Our conclusion is that the only model of interaction between a CAP dimer and DNA that adequately meets these three criteria requires that CAP binds to a left-handed, B-type DNA.

We have examined several possible modes of binding, and find only two ways in which a DNA molecule could interact with both small domains of a single CAP dimer with the approximate 2-fold axes of DNA and CAP coincident. Both ways involve the 22 Å-long F helices that are a dominant structural feature of the small domains, and are related by the approximate 2-fold axis between the small domains, are nearly parallel to each other and have their helix axes separated by about 34 Å. Furthermore, these helices are inclined at ~65–70° to the line connecting their centres.

The most complementary mode of interaction between the DNA and CAP is suggested by the 34 Å spacing and tilt of the 2-fold-related F α -helices. These two α -helices can fit snugly into two successive major grooves of left-handed B-DNA³⁵ (Fig. 4). In addition, the extended polypeptide chain is correctly positioned to interact with the sugar-phosphate backbone on either side of the two α -helices (Figs 4, 5c). Placing the α -helices in the major groove accounts nicely for the L8 and L29 mutations³⁶, methylation protection at the N⁷ of two pairs of guanines and the UV light-induced cross-linking of two symmetry-related thymines^{33,34}, because all these residues then interact with these helices (Fig. 5c). Furthermore, the phosphate moieties implicated by chemical modification to be interacting with CAP are almost all near the polypeptide backbone (Fig. 5c). In this way the protein can interact with at least 18–22 base pairs of DNA in a manner consistent with the footprinting data (refs 8, 34, 37 and A. Schmitz, personal communication) which show CAP protection spanning 25 base pairs.

Left-handed B-DNA has recently been shown to be a stereochemically acceptable conformation³⁵. Note that CAP cannot interact with left-handed Z-DNA^{38,39}, because the pitch is too large (45 Å) and there are no major grooves for the helices to fill. Furthermore, a purine–pyrimidine alternating sequence, requisite for formation of the Z-DNA helix, is not present in any known CAP recognition sites. Nevertheless, the known crystal

Fig. 4 Stereo drawing of the α -carbon backbone of the CAP dimer interacting with 22 base pairs of left-handed B-type DNA. One CAP subunit is drawn with dashed lines, the other subunit with solid lines. The coordinates for the DNA were derived from parameters given in ref. 37 and are presented in a space-filling representation using standard van der Waals radii. Note the extraordinary degree of general structural complementarity which extends over 18–20 base pairs (65–70 Å) of DNA.



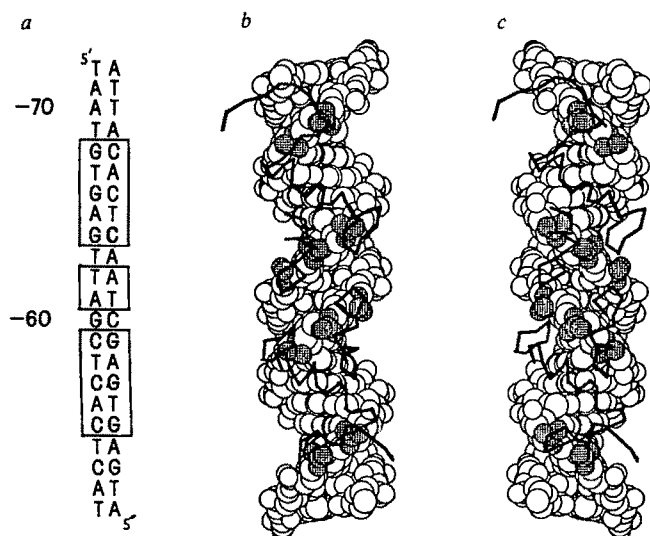


Fig. 5 *a*, Sequence of the CAP-binding region of the *lac* operon. Boxes enclose regions of 2-fold symmetry in the sequence. *b*, Space-filling model of right-handed B-DNA, with proximal regions of the protein α -carbon backbone superimposed. Contact sites predicted from chemical modification studies^{33,34} are shaded. *c*, The same for left-handed B-DNA.

structure of left-handed Z-DNA and the stereochemical possibility of left-handed B-DNA suggest that the energy difference between left- and right-handed DNA may not be large and that stabilization of a left-handed DNA by a protein might be possible.

In contrast, a complementary interaction between CAP and right-handed DNA is not possible (Fig. 6). In this case the F α -helices are skewed relative to the major groove and make contact only with a portion of the sugar-phosphate backbone. No protein can project into either groove to make specific contacts. This model does not account for the methylation protection, mutations or cross-linking observed with bases in the major groove.

Structural rearrangements of CAP which would allow it to bind with its F helices in the major groove of right-handed B-DNA are highly implausible. Rotation of the F helices by 60–70°, while constraining their centre-to-centre distance at 34 Å, would be required. Rotation of the small domains about the hinge points changes the centre-to-centre spacing of the F helices and hence is excluded. Rotation of the individual helices by 60–70° about their centre points would maintain the 34 Å spacing, but would require a major refolding of the small domains. As the crystals are grown in the conditions in which CAP binds specifically to DNA in solution, we expect that the observed crystal structure is in the active DNA-binding conformation. By analogy with virtually all other known protein

examples, there is no reason to expect that the structure of CAP in the crystals differs significantly from its structure in solution.

The only other model of DNA bound to CAP that we seriously considered was right-handed B-DNA placed in the shallow depression formed primarily by the two parallel F helices (model not shown). In this model the long axis of the DNA runs parallel to the F helices and makes minor contacts with the proximal ends of the D helices. Such a model seems less plausible because (1) the length of protein contact along the DNA helix would be only 35 Å or 10 base pairs, (2) contact of CAP along the DNA phosphate backbone would be along alternate sides of the helix, rather than along one face, and (3) there would be no structural complementarity between protein and either the major or minor groove of DNA.

Thus, we conclude that only left-handed DNA, bound to CAP as shown in Fig. 4, accommodates all the relevant data. At least two experimentally verifiable predictions are made by this model. First, the inclusion of amino acid side chains in the protein model should provide a DNA-binding site which is complementary to left-handed B-DNA in hydrogen bond donors and acceptors and charge interactions, as well as in overall shape as shown here. Second, specific binding of active CAP to appropriate CAP sites should reduce the winding number in closed circular DNA by about four, although nonspecific binding of either active or inactive CAP would probably not. Both experiments are in progress.

A possible mechanism by which cyclic AMP might exert its allosteric effect on specific CAP binding to DNA is suggested by the model of CAP-DNA (Fig. 4) and the two different orientations of the small domains observed in the two subunits. If the removal of cyclic AMP changes the orientation of the small domain of one subunit relative to its large domain, then the DNA-binding site would be destroyed. It would not be unreasonable for the two subunits to have the same structure (either both 'open' or both 'closed') in the absence of cyclic AMP and the dimer a 2-fold symmetry axis. A ligand-induced change in relative domain orientation has been demonstrated in hexokinase²⁵ and alcohol dehydrogenase²⁶.

Formation of left-hand B-DNA could activate transcription initiation by RNA polymerase

We suggest that CAP activates transcription by destabilizing, or melting, a right-handed DNA helix within the promoter region. The significantly different positions of the known CAP-binding sites relative to the promoter in the *lac*^{33,34}, *gal*³⁷ and *ara*⁸ operons rule out the possibility of a unique protein-protein interaction between CAP and RNA polymerase. Also, the observation that inhibition of DNA gyrase activity *in vivo* inhibits the expression of catabolite-sensitive genes⁴⁰, and the more recent *in vitro* observation that in the *lac* operon, CAP stimulation of polymerase activity is most effective when the

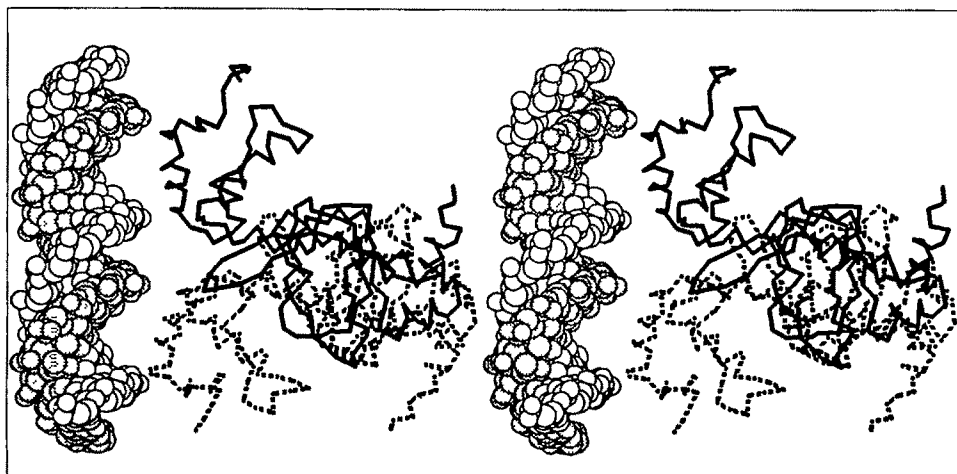


Fig. 6 Stereo drawing of a model of CAP and right-handed B-DNA⁴⁵ in approximately the same orientation as in Fig. 4. In this figure the DNA has to be moved further from the protein to avoid overlap of the F helices and the sugar-phosphate backbone. The contact surface between these macromolecules is strikingly less than in Fig. 4.

promoter and CAP site are in negatively supercoiled rather than relaxed linear DNA (W R McClure, personal communication), demonstrate that CAP is most effective when it is energetically favourable to unwind, and remove negative supercoils from a closed circular DNA helix⁴¹

Thus, although we do not know specific details of the transition, we propose that CAP traps specific regions of DNA in a left-handed helical conformation, and that the right-to-left helix transition denatures nearby regions of the helix in the promoter⁴², possibly the region specifically unwound by RNA polymerase itself⁴³

Anderson, Ohlendorf, Takeda and Matthews⁴⁴ have recently determined the course of the polypeptide backbone of λ cro protein and find that this repressor protein has two α -helices separated by 34 Å that can fit well into the major groove of right-handed B-DNA

Received 3 November 1980, accepted 24 February 1981

- 1 Zubay, G & Doty, P *J molec Biol* **1**, 1-20 (1959)
- 2 Sung, M T & Dixon, G H *Proc natn Acad Sci U.S.A.* **67**, 1616-1623 (1970)
- 3 Adler, K *et al Nature* **237**, 322-327 (1972)
- 4 Church, G M, Sussman, J L & Kim, S H *Proc natn Acad Sci U.S.A.* **74**, 1458-1462 (1977)
- 5 Zubay, G, Schwartz, D & Beckwith, J *Proc natn Acad Sci U.S.A.* **66**, 104-110 (1970)
- 6 Epstein, W, Rothman-Denes, L B & Hesse, J *Proc natn Acad Sci U.S.A.* **72**, 2300-2304 (1975)
- 7 deCrombrughe, B *et al Nature new Biol* **231**, 139-142 (1971)
- 8 Ogden, S, Haggerty, D, Stoner, C M, Kolodrubetz, D & Schleif, R *Proc natn Acad Sci U.S.A.* **77**, 3346-3350 (1980)
- 9 Musso, R E, DiLauro, R, Adhya, S & deCrombrughe, B *Cell* **12**, 847-854 (1977)
- 10 Anderson, W B, Schneider, A B, Emmer, M, Perlman, R L & Pastan, I *J biol Chem* **246**, 5929-5937 (1971)
- 11 Riggs, A D, Reiness, G & Zubay, G *Proc natn Acad Sci U.S.A.* **68**, 1222-1225 (1971)
- 12 Krakow, J G & Pastan, I *Proc natn Acad Sci U.S.A.* **70**, 2529-2533 (1973)
- 13 Eilen, E, Pampano, G & Krakow, J S *Biochemistry* **17**, 2469-2473 (1978)
- 14 Wu, C-W & Wu, F Y-H *Biochemistry* **13**, 2573-2578 (1974)
- 15 Wu, F Y-H, Nath, K & Wu, C-W *J molec Biol* **100**, 459-472 (1976)
- 16 Etien, E & Krakow, J S *J molec Biol* **114**, 47-60 (1977)
- 17 Kumar, S A, Murthy, N S & Krakow, J S *FEBS Lett* **109**, 121-124 (1980)
- 18 McKay, D B & Fried, M G *J molec Biol* **139**, 95-96 (1980)
- 19 Wyckoff, H W *et al J molec Biol* **27**, 563-578 (1967)
- 20 Rossmann, M G *Acta crystallogr A* **32**, 774-777 (1976)
- 21 Strahs, G & Kraut, J *J molec Biol* **35**, 503-512 (1968)
- 22 Harrison, S C, Olson, A J, Schutt, C E, Winkler, F K & Brucogne, G *Nature* **276**, 368-373 (1978)

These data suggest the following general model for proteins that regulate gene transcription (1) Two DNA-binding domains related by (approximate) 2-fold symmetry bind to one side of the DNA helix at the site of (approximate) 2-fold symmetry (2) Base-specific recognition can be accomplished by α -helices of the protein binding in the major groove of B-type DNA (3) Repressors, by binding to right-handed B-DNA, stabilize the right-handed helix and act as 'anti-melting' proteins, CAP protein, by binding to a left-handed helix, destabilizes the helix and facilitates binding and initiation by RNA polymerase

This work was supported by USPHS grant GM-22778, NSF grant PCM-77-25111 and by postdoctoral fellowships from the American Cancer Society and US Public Health Service to D B M We acknowledge assistance from Dr Harold W Wyckoff, Paul Pepin, Gerald Johnson, John Kennedy, Art Perlo, Steve Anderson, Michael Fried and Melody Lane

- 23 Abad-Zapatero, C *et al Nature* **286**, 33-39 (1980)
- 24 Bennett, W S Jr & Steritz, T A *Proc natn Acad Sci U.S.A.* **75**, 4848-4852 (1978)
- 25 Amzel, M & Poljak, R J *A Rev Biochem* **48**, 961-998 (1979)
- 26 Eklund, H, Samama, J-P, Wallen, L & Branden, C-I *J molec Biol* (in the press)
- 27 Takahashi, M, Blazy, B & Baudras, A *Biochemistry* **19**, 5124-5130 (1980)
- 28 Platt, T, Files, J G & Weber, K *J biol Chem* **248**, 110-121 (1973)
- 29 Geisler, N & Weber, K *Biochemistry* **16**, 938-943 (1977)
- 30 Jovin, T M, Geisler, N & Weber, K *Nature* **269**, 668-672 (1977)
- 31 Pabo, C O, Saver, R T, Sturtevant, J N & Ptashne, M *Proc natn Acad Sci U.S.A.* **76**, 1608-1612 (1979)
- 32 Ogata, R T & Gilbert, W *Proc natn Acad Sci U.S.A.* **75**, 5851-5854 (1978)
- 33 Simpson, R B *Nucleic Acids Res* **8**, 759-766 (1980)
- 34 Majors, J thesis, Harvard Univ (1979)
- 35 Gupta, G, Bansal, M & Sasisekharan, V *Proc natn Acad Sci U.S.A.* **77**, 6486-6490 (1980)
- 36 Dickson, R C, Abelson, J, Johnson, P, Reznikoff, W S & Barnes, W M *J molec Biol* **111**, 65-79 (1977)
- 37 Taniguchi, T, O'Neill, M & deCrombrughe, B *Proc natn Acad Sci U.S.A.* **76**, 5090-5094 (1979)
- 38 Wang, A H-J *et al Nature* **282**, 680-686 (1979)
- 39 Drew, H, Takano, T, Tanaka, S, Itakura, K & Dickerson, R E *Nature* **286**, 567-573 (1980)
- 40 Sanzey, B *J Bact* **138**, 40-47 (1979)
- 41 Cozzarelli, N R *Science* **207**, 953-960 (1980)
- 42 Dickson, R G, Abelson, J, Barnes, W M & Reznikoff, N S *Science* **187**, 27-35 (1975)
- 43 Siebenlist, U *Nature* **279**, 651-652 (1979)
- 44 Anderson, W F, Ohlendorf, D H, Takeda, Y & Matthews, B W *Nature* **290**, 754-757 (1981)
- 45 Arnott, S & Hukins, D W L *Biochem biophys Res Commun* **47**, 1504-1509 (1972)
- 46 Wilson, I A, Skehel, J J & Wiley, D C *Nature* **289**, 366-373 (1981)

Arrangements and rearrangements of sequences flanking the two types of rDNA insertion in *D. melanogaster*

Heli Roiha, J. Ross Miller, Lesley C. Woods & David M. Glover

Cancer Research Campaign, Eukaryotic Molecular Genetics Research Group, Department of Biochemistry, Imperial College of Science and Technology, London SW7 2AZ, UK

The two types of Drosophila melanogaster rDNA insertion occur at different sites in the 28S gene, separated by 51 nucleotide pairs. A short segment of rDNA is deleted at the site of the 5 kilobase type I insertion, whereas there is no such deletion of rDNA sequences which flank the type II insertion. Type I insertion elements in tandem arrays are flanked by very short segments of the 28S rRNA gene

A LARGE proportion of the ribosomal genes of *Drosophila melanogaster* contain an intervening sequence in the 28S rRNA gene¹⁻⁴. There are two types of intervening sequence which have been termed the type I and type II insertions. The type I insertions are found in about 60% of the rDNA units of the X chromosome⁵⁻⁷ and the type II insertions in about 15% of the rDNA units of both the X and Y chromosomes^{6,8-10}. Physical mapping studies showed that these two non-homologous types of insertion are apparently located within the rDNA at identical sites, suggesting that there might be a structural feature of this region which could promote either the acquisition or loss of these elements. A high proportion of type I sequences is also

present in the chromocentral heterochromatin^{7,11}, linked together in tandem arrays¹². As far as is known, the type II insertions are only found in ribosomal gene repeats.

Intervening sequence elements have subsequently been reported in the distal part of the gene for the large rRNA of several other organisms¹³⁻¹⁸. The nucleotide sequences at the junctions between the gene and the intervening sequence have been determined for the 28S gene from *Tetrahymena*¹⁹, the chloroplast rDNA from *Chlamydomonas*²⁰ and the mitochondrial rDNA of yeast²¹. In none of these cases is it possible to recognize a consensus sequence which might serve as a recognition site for enzymes which process the primary rRNA

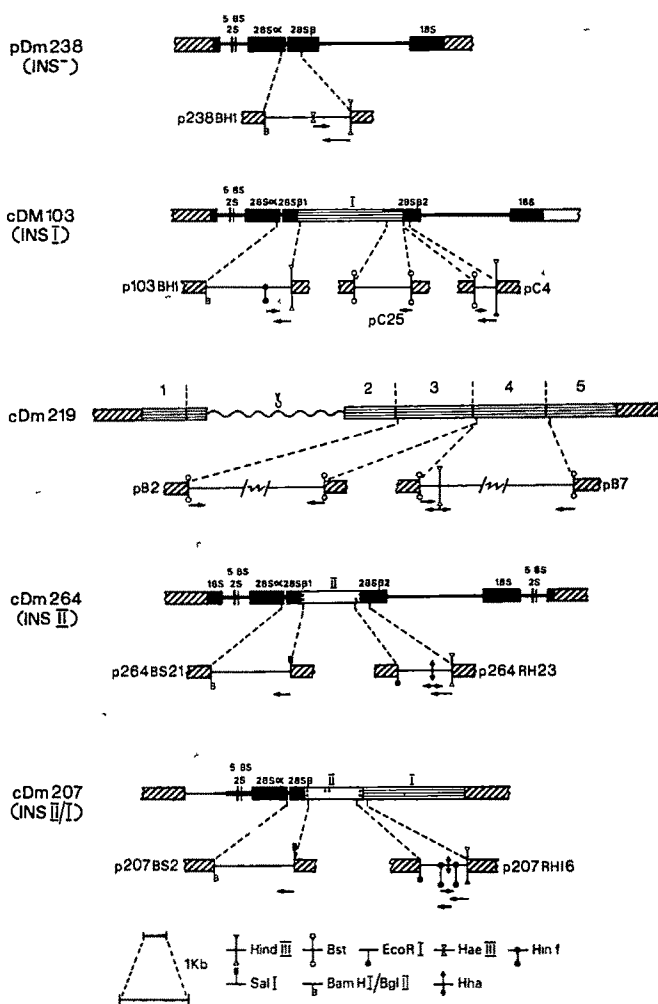


Fig 1 Sequencing strategies. The physical maps presented show cloned fragments of chromosomal DNA which contain representative arrangements of the rDNA insertion sequences. Subcloned fragments of these DNAs which contain the junction of the insertion sequences with their flanking sequences are shown beneath each map. The construction and isolation of the original genomic clones cDm103³¹, cDm219¹², cDm264¹⁰ and cDm207⁸ have been described elsewhere. The *D. melanogaster* DNA in pDm238 originates from the same set of randomly cloned segments as cDm219, cDm264 and cDm207, except that in this case a 12-kilobase *EcoRI* fragment from the original clone has been subcloned into the plasmid pBR322. All the subclones described were constructed in the appropriate sites of pBR322 as previously described¹². The exceptions are p238BH1 and p103BH1, in which the vector is pAT153³³, a non-mobilizable derivative of pBR322. The restriction site symbolized as *Bam*HI/*Bgl*II represents the product of ligating *Bam*HI and *Bgl*II termini in the vector and the *Drosophila* DNA, respectively. Only those restriction endonuclease cleavage sites used for either subcloning or sequencing are shown. The arrows beneath the maps of the subclones indicate the extent and direction of the sequences read from various restriction sites. All sequencing was performed using the techniques of Maxam and Gilbert³⁴. In each of the maps of the original clones, the ribosomal genes are indicated by filled blocks, the type I insertions by horizontal hatching and the type II insertion by dotted shading. These regions are not indicated in the subclones. The plasmid vectors are indicated by cross-hatching.

transcript, such as has been postulated for genes transcribed by RNA polymerase II²². Contrary to these cases, it is not certain whether those genes from *D. melanogaster* which contain the type I insertion sequence are transcribed into functional rRNA, as full-length transcripts corresponding to the gene and intervening sequence are extremely rare²³ and furthermore, there are rDNA units within the genome which do not contain intervening sequences.

We describe here the nucleotide sequences of the junctions between rDNA and the insertions and show that the two types of insertion occur at sites which are 51 nucleotides apart. The type I insertion occurs at an analogous site to the insertion in *Drosophila virilis* rDNA^{24,25}, but in all our clones we find that by comparison with the *D. virilis* sequence, 22 nucleotides of rDNA are deleted to the left-hand side of the insertion. There is no such deletion in the rDNA sequences flanking the type II

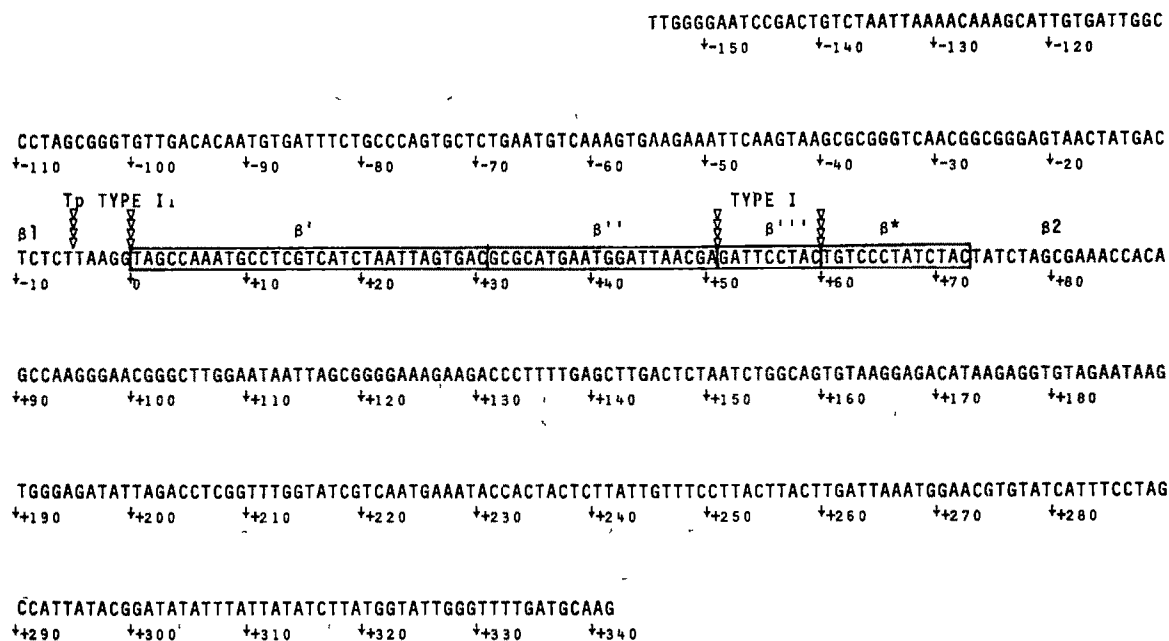


Fig 2 Sequence of a 497-nucleotide segment of the 28Sβ gene. This sequence was primarily obtained from p238BH1, a subclone of an uninterrupted unit of rDNA (see Fig 1). Plasmid DNA was labelled for sequencing either at the *Hin*dIII site by 'filling in' the cohesive terminus using reverse transcriptase or at the *Hae*III site using terminal transferase-mediated addition of an α -labelled ribonucleotide³⁴. The sequence of the equivalent regions in cDm103 and cDm264 have also been determined using comparable labelling techniques (see Figs 3 and 6 legends). Other than a short deletion of rDNA sequences, β'' , in cDm103 (see text) there are no differences in the segments of rDNA which have been sequenced in these plasmids. The two 'halves' of the 28Sβ gene on either side of the insertion are termed β 1 and β 2. We have arbitrarily begun our numbering from the position of the type II insertion which is marked with an arrow. In 28Sβ genes containing the type I sequence, the β'' segment is substituted by the insertion. The β^* sequence is analogous to a segment of rDNA which is duplicated in *D. virilis* rDNA units containing insertions²⁵, and which is also present in the clone cDm219. The β'' segment is also present in cDm219 whereas β' and adjoining β 1 sequences are not. The division of the 28Sβ gene into these units is to facilitate comparisons between various plasmids and does not necessarily denote any functional significance of these regions. The arrow marked Tp indicates the position of the intervening sequence in a homologous region of the 28S gene of *Tetrahymena pigmentosa*¹⁹.

a Left end

D. melanogaster (Dm103) β^{11}

GCGCATGAATGGATTAACGACGGACGTGTTTCGTT

D. virilis (DvL3-8) β^{11} β^{111} β^*

GCGCATGAATGGATTAACGAGATTCTCTAGTCCCTATCTAGTCTCAGTTTCGTTTCA

b Right end

D. melanogaster (Dm103) β^* β^2

TCGAAAAAGCATACTATGTCCTATCTAGTCTAGCGAAAC

D. virilis (DvL3-8) β^* β^2

AAGCCCACTGACTTTCAATGTCCTATCTAGTCTAGCGAAAC

Fig. 3 Comparison of the junctions of the type I insertion in *D. melanogaster* rDNA with those of the insertion in *D. virilis* rDNA. The *D. virilis* sequence shown is taken from Rae *et al.*²⁵ The sequences at the junctions of the type I insertion of cDm103 were obtained from the subclones pC4, pC25 and p103BH1 labelled by a 'filling-in reaction' with reverse transcriptase. The extent of the sequence determined for the Dm103 insertion in these experiments is indicated in Fig. 4. Segments of the 28S β gene are indicated above the boxed-in sequence and the insertion sequence is italicized.

insertion. The chromocentral type I insertion elements occurring in tandem arrays are separated by short DNA segments identical to parts of the 28S gene which flank the type I insertion in rDNA. We also analyse the structure of a segment of rDNA which has a 28S gene containing both types of insertion.

Location of insertions in the 28S rRNA gene

The 28S rRNA of *D. melanogaster* consists of four polynucleotide chains—the 2S RNA, 5S RNA, 28S α RNA and 28S β RNA—which are held together by hydrogen bonding. The relative positions of the DNA coding for these molecules are shown on the physical map of pDm238, a cloned rDNA unit which does not contain an insertion sequence (Fig. 1). The positions of type I and type II insertion sequences within the 28S β gene can be seen on the maps of the cloned segments cDm103^{1,2} and cDm 264¹⁰, respectively. The insertions divide the gene into two regions which have been termed 28S β 1 and 28S β 2 (Fig. 1). We have subcloned a restriction fragment from the uninterrupted unit pDm238 which contains the sites for the insertions, and restriction fragments from cDm103 and cDm264 which contain the junctions between the 28S β gene and the two insertion types. Figure 1 also shows restriction maps of these subcloned fragments and the DNA sequencing strategy which we followed to obtain a precise localization of the insertion elements within the gene.

The sequence of the 28S β gene in the vicinity of the insertions is given in Fig. 2. The sites of the insertion elements are evident by comparing this sequence with that of the comparable regions from cDm103 and cDm264 (see below). The two types of insertion occur at different sites within the 28S β gene, separated by 51 base pairs. The type II sequence is located at a site just five residues away from the position of the intervening sequence in the 28S gene in some strains of *Tetrahymena*¹⁹. The type I sequence is found at an analogous position to the insertion in the rDNA of *D. virilis*²⁵. The sequence of this same region in *D. virilis* is highly homologous and shows only a few single-base differences with the *D. melanogaster* rDNA.

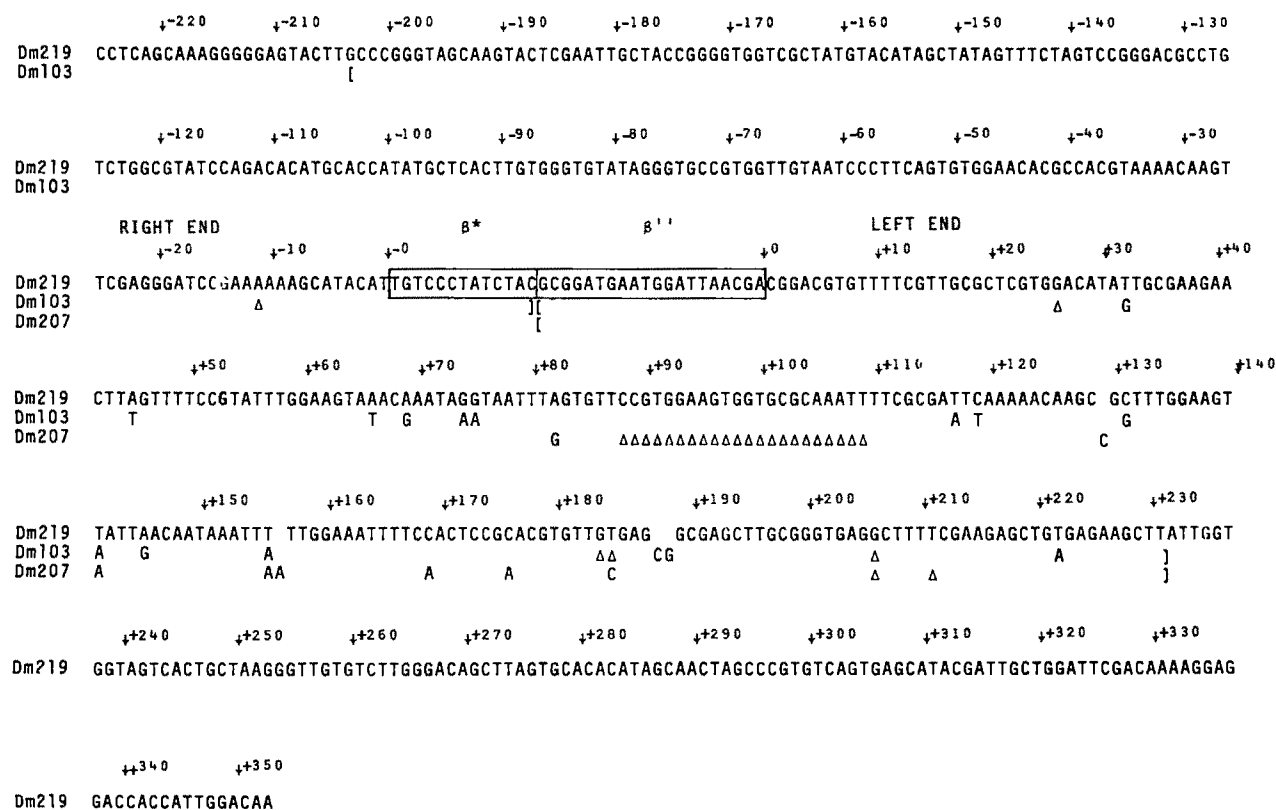


Fig. 4 Type I insertions. The continuous sequence from the junction of units 3 and 4 from cDm219 (Fig. 1). The two short segments of rDNA sequences β^* and β^{11} shown in boxes, occur at this junction (see text). Nucleotides 0 to -226 are from the right end of unit 3 and 0 to +337 are from the left end of unit 4. Sequencing was performed from the indicated *Hind*III and *Bst*I sites labelled at 5' termini using polynucleotide kinase and [γ -³²P]ATP (Fig. 1). A comparison is also given of type I sequences from cDm103 and cDm207 determined as indicated in the legends to Figs 3 and 6, respectively. Only the differences between these sequences and that of Dm219 are indicated. Δ Indicates a deleted nucleotide. The extent of the sequence determined for these clones is indicated by brackets. The right-hand end sequence of the Dm103 insertion comes from the subclones pC25 and pC4 and extends into the β^2 region of the 28S gene. The left-hand sequence of the cDm103 insertion is from p103BH1 and extends into the 28S β 1 region. The left-hand sequence of the cDm207 type I sequence is from p207RH16 (see Fig. 6).

Differences between the junctions of the major insertion in

D. melanogaster and *D. virilis*

Rae *et al.*²⁵ have demonstrated that in *D. virilis* rDNA a 14-nucleotide sequence present only once at the site of the insertion in the uninterrupted rDNA unit is repeated on both sides of the insertion. We find that the type I insertion in cDm103 is at this same site, flanked on its right-hand side by the same 14-nucleotide sequence as found in *D. virilis* which we have called β^* (Fig. 2). The structure of the left-hand side differs from that of *D. virilis*. There is no β^* sequence and, furthermore, there is a deletion of nine residues of rDNA which we have called the β''' segment (Fig. 3).

The arrangement of the β^* sequence in *D. virilis* resembles that found at the insertion sites of transposable elements of both prokaryotic²⁶ and eukaryotic^{27,28} organisms. The occurrence of an insertion at the same site in *D. melanogaster* suggests that if the insertions arose in the rDNA of ancestral Dipteran species as a result of transposition, then the resulting unit has been reasonably stable throughout evolution. Alternatively, the duplication of the β^* sequence may be analogous to the duplication of *att* site sequences resulting from homologous recombination which occurs between the genomes of *Escherichia coli* and its lambdoid phages²⁹. In this case site-specific recombination would perhaps be a continuous process in the rDNA of *Drosophila* species. The deletion of rDNA sequences from the left-hand side of the *D. melanogaster* insertion is also analogous to the type of deletions that can be generated on excision of either prokaryotic transposable elements or lambdoid prophage. Although rDNA units containing such deletions are clearly stably maintained to some extent within the genome it remains of interest to know whether there are rDNA units within *D. melanogaster* which have the duplicated segments found in *D. virilis*. This would help answer some of the above questions.

Small rDNA segments occur in the chromocentral type I sequences

As many as 50% of the type I sequences are not inserted into rDNA but are located within the chromocentral heterochromatin and also at a single site on chromosome 4 in band 102C^{7,12}. The cloned DNA segment cDm219 originates from the chromocentral heterochromatin and contains five discrete units of the type I sequence arranged in tandem. We have determined the sequences between the junctions of the units of cDm219 following the strategy outlined in Fig. 1. The sequence at the junction of units 3 and 4 of the tandem array is given in Fig. 4, where the sequences corresponding to the left- and the right-hand ends of the rDNA insertion have positive and negative numbers, respectively. At the left hand of these units we find

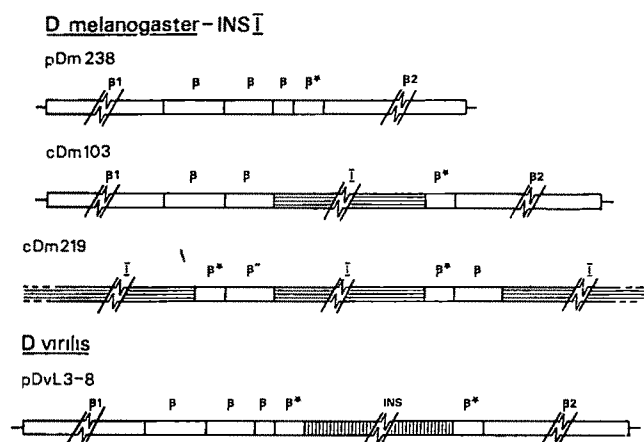


Fig. 5 A schematic comparison of the segments of the 28S β gene which flank the type I insertions in cDm103 and cDm219 with those flanking the insertion in *D. virilis* rDNA.

20 nucleotides of rDNA, the β'' segment, the same segment of the 28S β gene which flanks the left hand of the type I insertion in rDNA (Fig. 3). At the right-hand end of these units the β^* segment flanks the insertion as in Dm103 (Fig. 3), and at the junctions between the tandemly arranged units of cDm219 the β'' segment is linked to the β^* segment. The relative positions of these small segments of the 28S β gene which flank the type I insertion of *D. melanogaster* rDNA and the insertion in *D. virilis* rDNA are shown schematically in Fig. 5. The finding of short segments of rDNA which flank units of chromocentral type I sequences strongly suggests that type I elements can be transposed from the nucleolus and that co-transposition of adjoining rDNA segments can occur during this process. The excised unit could then be amplified either by autonomous replication or by unequal exchange following integration into its new chromosomal site. If the insertion elements of both *D. virilis* and *D. melanogaster* are indeed mobile, these observations would suggest that only the right-hand β^* element is required for transposition. Perhaps it is significant in this respect that within the rDNA sequence to the right of β^* in both *D. virilis* and *D. melanogaster* there is a direct duplication of part of the β^* sequence (residues 66–79, Fig. 2).

The continuous sequence given in Fig. 4 is that of the junctions of the cDm219 subclones pB2 and pB7. We have also sequenced about 200 residues on either side of the junctions in other subcloned type I units of cDm219, and have found no differences between these junctions and the sequence shown in Fig. 4. Figure 4 also compares the sequences from the left and right ends of the rDNA insertion in cDm103. Within this limited sequence there seems to be a significant difference in the degree of conservation between the left and right ends of these insertion elements. The only base difference within a comparable 226-nucleotide sequence from the right-hand end is the presence of an extra dA residue in cDm219. At the left end, however, there are 19 nucleotides which differ within a comparable 231-nucleotide sequence. There is similar variability in an equivalent sequence from the left-hand end of the type I insertion found in cDm207 (see below), this sequence has 10 single-base differences when compared with this Dm219 element and 20 compared with cDm103. There is also a 22-nucleotide deletion in this region of cDm207. In this respect, we note that although there is considerable length variation of the type I sequences, it is the sequences from the right-hand side of the unit which are preserved^{3,6}. It will be interesting to examine the nature of junctions of these shorter type I elements with adjoining rDNA.

Junctions of the type II insertions

About 15% of the *D. melanogaster* rDNA units have insertions with cleavage sites for *EcoRI*^{6,8}. We have characterized several cloned segments of rDNA containing complete type II sequences and find no homology with the type I sequences¹⁰. We chose to determine the sequence at the junctions of the type II insertion of cDm264 because this insertion has the most commonly observed length and restriction endonuclease cleavage pattern. We find that there is a direct sequence repeat at the left-hand junction of the type II insertion with the 28S β gene. The repeated sequences, GGGGAG, are contiguous and are formed partly from rDNA sequence and partly from insertion sequence (Fig. 6a). This is quite distinct from the duplication of the β^* sequence flanking the *D. virilis* insertion. The right-hand junction of the type II insertion shows a striking tract of 22 dA residues, which are followed by the β' , β'' , β''' , β^* and β^2 segments of the 28S gene (Fig. 6b, c). Unlike the case of rDNA containing the type I insertion, there is no deletion of adjacent rDNA sequences.

The deletion in Dm103 makes it extremely unlikely that units of this type are ever used in the production of mature rRNA. This is in accord with the findings that large nuclear transcripts containing type I sequences are extremely rare, and that the most abundant transcript is a cytoplasmic molecule of 1 kilobase, present in about 50 copies per cell²³. Choori³⁰, however, has described the presence of long transcription units in actively

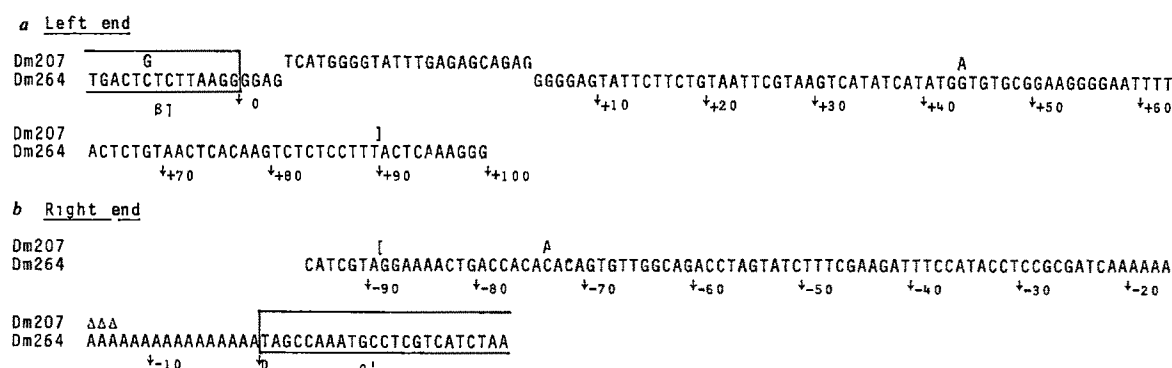


Fig 6 Sequences at the junctions of the type II insertions within rDNA. *a*, The sequence of the left end of the type II sequences in cDm264 and cDm207 comes from the subclones p264BS21 and p207BS2, respectively. The restriction fragments were labelled at the *Sall* site by 'filling-in' the *Sall* cohesive end using reverse transcriptase. Only the regions in which the cDm207 sequence differs from the cDm264 sequence are shown. *b*, The sequence of the right end of the type II sequence in cDm264 and cDm207 from the subclones p264RH23 and p207RH16, respectively. *Hind*III and *Hinf* termini were 3'-labelled using reverse transcriptase and the *Hae*III termini 3'-labelled using terminal transferase. The single bases of the cDm207 sequence which differ from cDm264 are indicated. *c*, Schematic diagram of how these rDNA segments differ to the right of the type II insertion. The segments β' and β'' , defined in Fig 2, are found in each plasmid. In cDm264, β' is contiguous with the remainder of the 28S β gene, whereas in cDm207 it is linked to the left end of the type I sequence. This type I sequence in cDm207 differs from those of cDm219 and cDm103 at the sites indicated in Fig 4.

transcribed chromatin which could represent transcription of rDNA units containing insertions. The distribution of these long transcription units correlates with the distribution of type II sequences in the rDNA of the X and Y chromosomes. It is possible, therefore, that rDNA containing type II insertions has a similar transcriptional pattern to that found in *Tetrahymena* rDNA where the insertion sequence is present in the primary transcript and is subsequently removed. Moreover an A+T-rich region seems to be a common feature of the right-hand junction of the insertions in the rDNA of *Tetrahymena*¹⁴, *Chlamydomonas* chloroplasts²⁰ and yeast mitochondria²¹, for which the A+T content of the rightmost 25 nucleotide pairs is 72, 80 and 76%, respectively. Indeed, it has been postulated that A+T-rich regions might be an important feature of other processing sites in ribosomal transcripts.

A 28S β rRNA gene containing both insertions

We have previously found that the clone cDm207 contained both types of insertion and that within our mapping resolution the junction of type II with type I sequences corresponded to the right and left ends of these respective sequences, as they are found singly within rDNA¹⁰. Nucleotide sequencing indicates that the type II element occurs at exactly the same site as in cDm264 (Fig 6*a, b*). Interestingly, the type II sequence in cDm207 differs from that of cDm264 by an additional 23 nucleotides which occur between the two repeated hexameric sequences. A tract of 19 dA residues occurs at the right-hand end of the type II insertion in cDm207. However, it is linked to only 51 nucleotides of rDNA, corresponding to the β' and β'' segments, which intervene between the sites of the two insertions. The junction between the β' segment and the type I insertion is identical to that found in cDm103 and in the units of cDm219. The comparative structures of cDm207 and cDm264 are indicated schematically in Fig 6*d*.

We previously noted that the restriction map of the type I insertion in cDm207 is more similar to that of a Dm219 unit than of the Dm103 rDNA insertion. A sequence comparison within a limited segment of these DNAs is consistent with this observation (see above and Fig 4). The chromosomal segment Dm207 also contains an unusual interruption of the internal transcribed spacer region (see Fig 1), and this has previously led us to propose that it was the product of a recombinational event between an rDNA unit containing type II sequences and type I sequences elsewhere in the genome. That recombinational events occur to produce rDNA segments containing both types

of insertion is to be expected because the insertions occur at different sites. Several such clones have been isolated by ourselves and others and we expect that sequencing will show such clones to have analogous organizational features to those which occur between the two types of insertion in Dm207.

This work was supported by project grants from the MRC. D M G holds a career development award from the Cancer Research Campaign. We thank Simon Kidd for his interest in this work, for providing us with data on the positioning of restriction endonuclease cleavage sites and for the construction of several of the subcloned plasmids described here. We also thank Therese Anderton for help with sequencing techniques and Sue Hayman for typing the manuscript, especially the figures.

Note added in proof We have recently sequenced the junctions between rDNA and a 1.0 kilobase type I insertion. This insertion is flanked by duplicated β^* segments and there is no deletion of rDNA sequences.

Received 28 November 1980, accepted 5 February 1981

- Glover, D M & Hogness, D S *Cell* **10**, 167-176 (1977)
- White, R L & Hogness, D S *Cell* **10**, 177-192 (1977)
- Wellauer, P K & Dawid, I B *Cell* **10**, 193-212 (1977)
- Pellegrini, M, Manning, J & Davidson, N *Cell* **10**, 213-224 (1977)
- Tartof, K D & Dawid, I B *Nature* **263**, 27-30 (1976)
- Wellauer, P K, Dawid, I B & Tartof, K D *Cell* **14**, 269-278 (1978)
- Peacock, W J, Appels, R, Endow, S & Glover, D M *Genet Res* **38**, (1980)
- Glover, D M *Proc natn Acad Sci U.S.A.* **74**, 4932-4936 (1977)
- Dawid, I B, Wellauer, P K & Long E O *J molec Biol* **126**, 749-768 (1978)
- Roiha, H & Glover, D M *J molec Biol* **140**, 341-355 (1980)
- Dawid, I B & Botchan, P *Proc natn Acad Sci U.S.A.* **74**, 4233-4237 (1977)
- Kidd, S J & Glover, D M *Cell* **19**, 103-119 (1980)
- Wild, M A & Gall, J G *Cell* **16**, 565-573 (1979)
- Dim, N & Engberg, J *J molec Biol* **134**, 555-574 (1979)
- Gubler, U, Wyler, T & Braun, R *FEBS Lett* **100**, 347-350 (1979)
- Campbell, G R, Littau, V C, Meiera, P N, Allfrey, V G & Johnson, E M *Nucleic Acids Res* **6**, 1433-1447 (1979)
- Devries, H, Dejonge, J C, Bakker, M, Meurs, H & Kroon, A *Nucleic Acids Res* **6**, 1791-1803 (1979)
- Hahn, U, Lazarus, C M, Lunsdorf, H & Kuntzel, H *Cell* **17**, 583-595 (1979)
- Wild, M A & Sommer, M *Nature* **283**, 693-694 (1980)
- Allet, B & Rachaux, J *Cell* **18**, 55-60 (1979)
- Dujon, B *Cell* **20**, 185-197 (1980)
- Breathnach, R, Benoist, C, O'Hare, K, Gannon, F & Chambon, P *Proc natn Acad Sci U.S.A.* **75**, 4853-4857 (1978)
- Long, E O & Dawid, I B *Cell* **18**, 1185-1196 (1979)
- Barnett, T & Rae, P M *M. Cell* **16**, 763-775 (1979)
- Rae, P M M, Kohorn, B D & Wade, R P *Nucleic Acids Res* **8**, 3491-3504 (1980)
- Calos, M P & Miller, J H *Cell* **20**, 579-595 (1980)
- Dunsin, P, Brorein, W J, Simon, M A & Rubin, G M *Cell* **21**, 575-580 (1980)
- Farabaugh, P J & Fink, G R *Nature* **286**, 352-356 (1980)
- Landy, A & Ross, W *Science* **197**, 1147-1160 (1977)
- Chooi, W Y *Chromosoma* **74**, 57-74 (1979)
- Glover, D M, White, R L, Finnegan, D J & Hogness, D S *Cell* **5**, 149-157 (1975)
- Grunstein, M & Hogness, D S *Proc natn Acad Sci U.S.A.* **72**, 3961-3965 (1975)
- Twigg, A J & Sherratt, D *Nature* **283**, 216-218 (1980)
- Maxam, A M & Gilbert, W *Meth Enzym* **65**, 499-560 (1980)

Structure of the cro repressor from bacteriophage λ and its interaction with DNA

W. F. Anderson^{*†}, D. H. Ohlendorf^{*}, Y. Takeda[‡] & B. W. Matthews^{*}

^{*} Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA

[†] MRC Group on Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

[‡] Chemistry Department, University of Maryland, Baltimore County, Catonsville, Maryland 21228, USA

The three-dimensional structure of the 66-amino acid cro repressor protein of bacteriophage λ suggests how it binds to its operator DNA. We propose that a dimer of cro protein is bound to the B-form of DNA with the 2-fold axis of the dimer coincident with the 2-fold axis of DNA. A pair of 2-fold-related α -helices of the repressor, lying within successive major grooves of the DNA, seem to be a major determinant in recognition and binding. In addition, the C-terminal residues of the protein, some of which are disordered in the absence of DNA, appear to contribute to the binding.

INTERACTIONS between proteins and nucleic acids are of central importance in molecular biology. Some proteins bind single-stranded polynucleotides while others recognize double-stranded ones. None of these interactions is well understood in molecular terms. In particular, the way in which proteins recognize specific base sequences within double-stranded DNA has remained enigmatic.

The cro protein from bacteriophage λ (hereafter called 'cro') is a repressor molecule which recognizes and binds tightly certain sequence-specific operator regions within the DNA of the bacteriophage. Cro is the smallest repressor protein that has been characterized, and is of interest both because of its role in the delicately balanced regulatory system of bacteriophage λ , and as a general model for protein-DNA interaction.

We describe here the three-dimensional structure of the cro repressor and its apparent mode of interaction with operator DNA. The postulated interaction between repressor and operator is consistent with a variety of chemical and genetic evidence, and may provide a general model for the interaction of proteins with helical DNA.

The bacteriophage λ genome encodes two repressor proteins that play essential, contrasting parts in regulating the development of the phage. The product of the cI gene, the cI or λ repressor ('cI'), is required for the maintenance of the lysogenic state, whereas the product of the cro gene is necessary during the lytic cycle of phage development (for reviews see refs 1-4, 33, 34). Both the cro and the cI repressors compete for and bind two operator regions within the phage genome. These operator regions each consist of three binding sites for the repressor molecules, and the respective repressors have different relative affinities for each of the sites within the operator regions^{5,6}. The use of these binding sites by cro and cI provides for a switch between the lysogenic and lytic pathway. The cro/cI interaction is the clearest example of such a developmental switch³³ and the structural results presented here begin an understanding of this switch at the molecular level.

The cro protein is well characterized^{7,8}. It is a small basic protein 66-amino acid residues long, with a monomer molecular weight of 7,351. Both the amino acid sequence of the protein⁹ and the nucleotide sequence of the cro gene have been determined¹⁰.

Structure determination

The structure was determined by the method of isomorphous replacement¹¹⁻¹³. Protein purification and crystallization from 1.1 to 1.3 M phosphate were as described previously^{8,14}. The space group is R32 and the cell dimensions in the hexagonal system $a = b = 91.9 \text{ \AA}$, $c = 268.9 \text{ \AA}$.

Potential heavy-atom derivatives were initially screened by precession photography, and, later, by using a diffractometer to collect 5- $\text{\AA}$ -resolution data sets. The first heavy-atom derivative to be interpreted was obtained using PtCl_4^{2-} and was initially solved by a 'brute force' approach in which all possible indi-

vidual sites, and then combinations of sites, were explored (ref 15 and unpublished results of B W M). Using the phases obtained from this platinum derivative, other potential derivatives were then evaluated in the normal way^{12,13}. In the end, five derivatives were selected for high-resolution data collection (Table 1). All the derivatives have multiple binding sites, as expected with multiple copies of the protein present in the crystallographic asymmetric unit (see below). Higher-resolution data were collected by oscillation photography¹⁶ which permitted collection of data sets to a nominal resolution of 2.8 $\text{\AA}$ or better from a single crystal. The crystals often grow as long triangular prisms¹⁴, in which case they can be translated between series of exposures. Some data collection statistics are included in Table 1.

Two preliminary electron density maps to 5 $\text{\AA}$ resolution were calculated, one using the data collected by diffractometry, and the other using a subset of the higher-resolution oscillation data. The two maps were quite similar, and, as discussed below, were important in working out the arrangement of the cro monomers within the crystallographic asymmetric unit.

Molecular symmetry

The initial analysis of the cro protein crystals showed them to contain multiple copies of the protein in the asymmetric unit. On the basis of the density measurement of two cross-linked crystals, there seemed to be six monomers, that is, three dimers, per asymmetric unit¹⁴. However, examination of the 5- $\text{\AA}$ -resolution electron density map revealed that there were four monomers per asymmetric unit, not six. The error in the previous estimate could be due either to the presence of the cross-linking agent or to errors in density due to the limited amount of material available.

We attempted to determine the relation between the monomers in two ways—by inspection of the 5- $\text{\AA}$ -resolution electron density maps, and by consideration of the relation between the sites of heavy-atom binding. Initial inspection of the 5- $\text{\AA}$ -resolution map revealed an obvious 2-fold symmetry.

Table 1 Heavy-atom derivatives

Derivative	No of sites	R_{Merge} (%)	R_{Der} (%)	Reflections above 3σ	R_{wc} (%)	$\langle f_{\text{H}} \rangle$	E
Native	—	9.5	—	9184	—	—	—
PtCl_4^{2-}	6	9.6	14.5	5068	30.5	64.3	54.5
Hg^{2+}	16	8.9	12.9	3616	44.9	113.4	72.4
Au^{2+}	16	8.7	14.0	7273	47.1	59.5	46.8
EMTS	12	0.3	11.0	7856	48.7	39.0	32.2
Mersalyl	13	10.2	14.0	5985	46.1	62.5	51.5

Overall figure of merit¹⁸ to 2.8 $\text{\AA}$ resolution, 0.45. $R_{\text{Merge}} = 100 \sum (I_i - \langle I \rangle) / \sum \langle I \rangle$, $R_{\text{Der}} = 200 \sum |I_{\text{PH}} - I_P| / \sum (I_{\text{PH}} + I_P)$, $R_{\text{wc}} = 100 \sum \omega |F_{\text{PH}} - F_P| - f_{\text{H}}| / \sum \omega |F_{\text{PH}} - F_P|$ (f_{H} is the mean heavy-atom scattering, E the average lack of closure for centric reflections¹⁸). EMTS, ethyl mercury thiosalicylic acid. Hg^{2+} data set collected by precession photography.

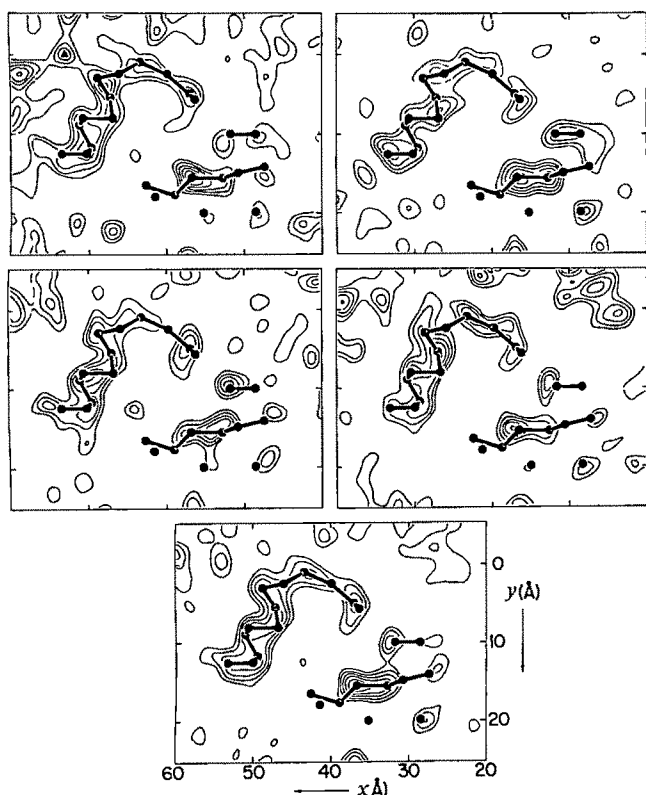


Fig. 1 Corresponding electron density sections through each of the four cro molecules in the crystallographic asymmetric unit, transformed into a common coordinate frame. Clockwise from top left the density is for molecules O, A, B and C, with the averaged density below α -carbon atoms within ± 2.5 Å of the section are indicated by the solid circles

axis lying approximately in the plane $z = 1/12$. This 2-fold axis also related, in pairs, most of the major heavy-atom sites. Further analysis of the 5-Å map, facilitated by the use of an MMS-X computer graphics system¹⁷, suggested that there were two additional 2-fold axes inclined to the first. The positions of these local symmetry elements were confirmed by calculating correlation coefficients between the electron density of the individual monomers, rotated and translated to superimpose one monomer on another (Table 2). The four monomers are related in pairs by mutually perpendicular intersecting (or nearly intersecting) 2-fold axes. In other words, the four monomers are arranged in the crystal with local 222-point symmetry. The symmetry of the molecules in the crystal does not itself imply anything about the state of aggregation in solution. In principle, the protein could exist as individual monomers, as dimers with 2-fold symmetry, or as tetramers. However, the specific interactions which we observed between the four monomers (see below) suggests that cro may exist in solution as a tetramer. The extent to which the molecular arrangement may depart from exact 222 symmetry, and the potential biological relevance of such deviations from exact symmetry, remain to be determined.

Most, but not all, of the heavy-atom sites are related to each other by the 222 local symmetry elements. For each 2-fold axis, related pairs of sites can be superimposed within ~ 1.9 Å (Table 2). Although the 2-fold symmetry was used as an additional check in the interpretation of the heavy-atom derivatives, it was not imposed on the heavy-atom coordinates.

The electron density map

An electron density map was calculated to a nominal resolution of 2.8 Å using the five heavy-atom derivatives described in Table 1 (ref. 18). Although the phase angles are well determined to about 3.5 Å resolution, their accuracy at higher angles is limited by the scattering intensity of the derivative crystals, which becomes very weak at higher resolution, and also, in part, by non-isomorphism.

To improve the accuracy of the electron density map and facilitate its initial interpretation, the density of the four monomers in the asymmetric unit was averaged (Table 2, Fig. 1). If the assumed transformations are correct and all four monomers are, in fact, identical, the averaging process should improve the accuracy of the electron density map. In practice, the averaging process clearly improved the overall 'cleanliness' of the map, but also resulted in a loss of detail and reduction in the effective resolution (Fig. 1). Presumably, application of more sophisticated averaging techniques¹⁹ and refinement of the individual monomers will lead to a more precise description of the structure.

The averaged electron density map was displayed both as a 'mini-map' and on a scale of 2 cm = 1 Å in an optical comparator^{20,21}. Several helices and a region of β -sheet were immediately apparent, and the course of the polypeptide backbone could be readily followed from one end to the other. The direction of the polypeptide chain was initially deduced from the observation that the major platinum binding site was at a position 12 residues from one terminus, obviously corresponding to Met 12. From this point the known amino acid sequence of the protein^{9,10} could be compared with the electron density map. Although the interpretation of the electron density map is fully consistent with the sequence, note that the limited resolution of the current map, together with the averaging process described above, results in poorly defined density for the respective side chains. The backbone density is strong and unambiguous from the amino terminus to residue Lys 62, but at this point the density becomes very weak, indicating that the four carboxy-terminal residues are disordered in the crystals. This appears to be the case for each of the four cro monomers.

Conformation of the cro repressor

The conformation of a cro monomer is illustrated in Fig. 2. The structure is very simple, consisting of three strands of antiparallel β -sheet (residues 2–6, 39–45 and 48–55), and three α -helices (residues 7–14, 15–23 and 27–36). As described below, the C-terminal residues form an extended 'arm' (Fig. 2) which interacts with another monomer and seems to participate in DNA binding.

The detailed arrangement of the four repressor monomers in the crystals is shown in Fig. 7. The respective local 2-fold symmetry axes relating the monomers are designated P, Q and R, and the monomers themselves O ('original'), A, B and C (Table 2). It has been assumed for some time that the cro

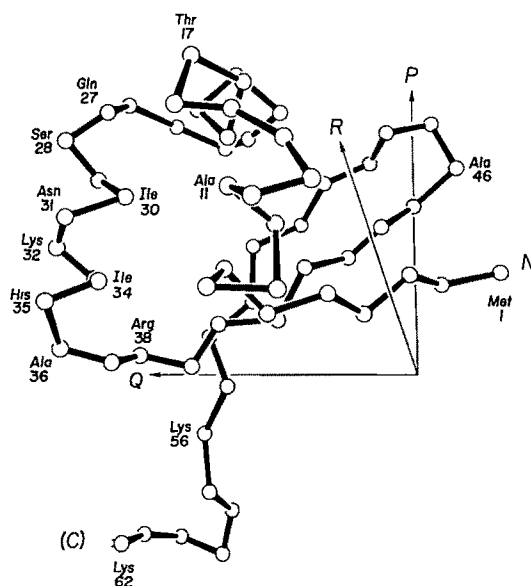


Fig. 2 Backbone of monomer O of the cro repressor protein. As discussed in the text, the four C-terminal residues 63–66 are disordered. The local orthogonal 2-fold symmetry axes P, Q and R relate the molecule shown to monomers A, B and C, respectively.

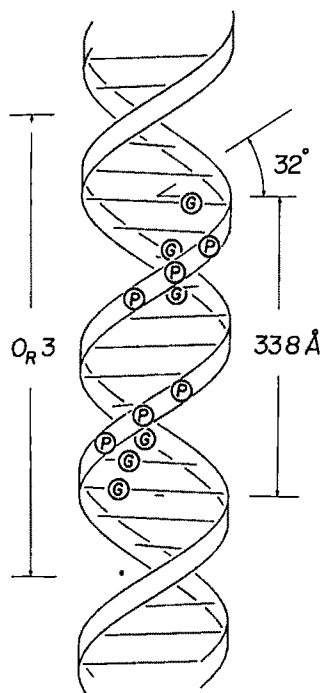


Fig. 3 DNA operator O_{R3} , drawn in the B form (after Ptashne *et al.*⁴) Ethylation of the phosphates, labelled P, hinders cro repressor binding, binding of cro protects the N^7 positions of guanines, labelled G, from methylation by dimethyl sulphate. The figure is based on experiments of Johnson, Ptashne and collaborators.^{4,5,35}

repressor exists as a dimer in physiological conditions⁸. Therefore, we expected the four cro monomers in the crystal to consist of two dimers. However, inspection of the molecular packing does not obviously support this assumption. The monomers O and B (and, similarly A and C) are related by the Q axis, and have extensive interactions between their C-terminal arms, each of which lies against, and is, in part, hydrogen bonded to, its 2-fold-related partner. It seems likely that this interaction persists in solution. However, there are also specific interactions about the R symmetry axis, where the polypeptide backbone of one monomer hydrogen bonds to the 2-fold-related backbone of the other monomer so as to expand the three-strand β -sheet of the monomer into a six-strand antiparallel sheet in the dimer. In addition, the amino terminus of monomer O is linked by an ionic interaction to Glu 53 of monomer A, with similar interactions for the other amino termini. The four amino-terminal methionines are grouped together and participate in a distinct hydrophobic region at the centre of the cro tetramer. Thus, although the packing arrangement seen in the crystal does not prove that cro exists as a tetramer in solution, it does suggest this. Whether cro exists as a dimer or tetramer has no bearing on the proposed model for DNA binding.

Interaction of cro with DNA

Both cI and cro repressors compete for three sites within the left and the right operators of bacteriophage λ . In the right operator there are three binding sites (O_{R1} , O_{R2} and O_{R3}) each consisting of 17-base pair regions separated by spacers of 6 or 7 base pairs²². The sequences of the three 17-base pair regions are similar to each other, but not identical, and, in addition, the sequence within each region is approximately symmetric. Mutations within the 17-base pair regions reduce the binding affinity of both cro and cI, whereas mutations in the spacer regions have no effect on binding (refs 4, 6, 23, 24 and refs therein). Chemical probe experiments^{4,5,35} indicate that both cro and cI repressors bind primarily along one face of the DNA double helix and protect many of the same groups. The results of such experiments for cro binding to O_{R3} are illustrated in Fig. 3. Methylation by dimethyl sulphate of the N^7 of six guanines in the major groove is prevented by the presence of cro, but

methylation of N^3 of adenine, exposed in the minor groove of the DNA, is not prevented. Furthermore, ethylation of any of the six phosphates shown in Fig. 3 interferes with cro binding⁴. The pattern for cI is very similar, but is slightly more extended, and includes four additional phosphates. Such experiments suggest that both cro and cI contact the DNA in the major, but not the minor, groove^{4,5}. Furthermore, it seems likely that both repressors bind to the DNA with two symmetry-related subunits of the protein in contact with the two symmetry-related halves of the operator, and the 2-fold axis of the protein coincident with the 2-fold axis of the DNA. In this way the complex of bound repressor and operator would have a common 2-fold symmetry axis. There would, of course, be slight deviations from exact 2-fold symmetry because of the inexact symmetry of the DNA base sequence, and also, possibly, due to the protein.

To determine possible modes of binding of cro to DNA, we looked for a feature of the cro monomer which, together with its 2-fold related mate, would form a structure complementary to that of the operator DNA, and which, when aligned on the DNA, would be consistent with the protection experiments summarized in Fig. 3. An obvious candidate is the helix from residues Gln 27 to Ala 36. This helix lies along the surface of the protein, with isoleucines 30 and 34 contributing to the hydrophobic interior of the monomer, but with the helix otherwise protruding from the surface of the protein (Fig. 2). The corresponding helix in monomer B, related by rotation about axis Q (Fig. 4) is parallel, and the respective helices have a centre-to-centre distance of 34 Å, the same as the separation between successive major grooves of B-form DNA (Fig. 3). Furthermore, the angle made between the axes of the two α -helices and the line connecting their midpoints is equal to the angle between the major groove of the DNA and its long axis (Figs 3, 4). The sense of the respective angles is such that the two helices can be accommodated, in a very natural way, within successive major grooves of DNA in its normal Watson-Crick B conformation^{25,26}.

This hypothetical arrangement was checked by fitting a model of the cro molecule to that of DNA. We required that the 2-fold axis relating the cro monomers had to coincide with the 2-fold axis through the centre of the DNA operator. Thus, the fitting of cro to the DNA is very restrictive, allowing only two degrees of freedom—one a rotation of cro relative to the DNA, about the common symmetry axis; the second a translation of cro along the symmetry axis, which varies the distance between cro and the long axis of the DNA. The results of this first, albeit crude, fitting process are shown in Figs 5, 6 and 7.

The complementarity between protein and DNA is striking. In addition to the overall correspondence between the respective backbones of the DNA and protein, residues Ser 28, Ala 29, Asn 31, Lys 32, Ala 33 and His 35 of the above-mentioned α -helix are in a position to interact either with the exposed hydrogen bonding groups of the major groove, conferring specificity, or with the negatively charged phosphates, promoting generalized binding. Nearby residues, including Arg 38 and Lys 39, are also suitably located to contribute to binding and recognition. Close to the 2-fold symmetry axis, the sequences Glu 54–Val 55–Lys 56 of the two monomers form a pair of antiparallel β -sheet strands parallel to the minor groove which possibly interact with the DNA in the manner proposed

Table 2 Relations between the four cro monomers

Symmetry axis	Pairs of heavy-atom sites	r m s error in heavy-atom sites	Monomers related	Electron density correlation
P	26	1.80	O,A	0.58
Q	26	1.97	O,B	0.54
R	25	1.94	O,C	0.55

$C_{ij} = \sum \rho_i \rho_j / (\sum \rho_i^2 \sum \rho_j^2)^{1/2}$ where the summations are over all values, ρ_i , of the 2.8 Å-resolution electron density map within 15 Å of the centre of the respective monomers.

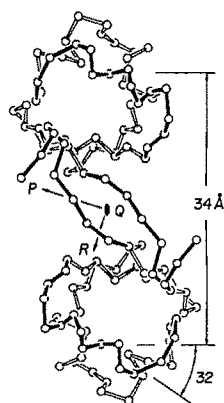


Fig. 4 Dimer of two cro molecules viewed along the Q symmetry axis. The P and R symmetry axes are also shown. Regions of the protein backbone closest to the viewer, and presumed to interact with DNA, are drawn solid. These regions include two α -helices, 34 Å apart, inclined at 32°, and also a pair of extended strands close to the Q-axis.

by Church *et al.*²⁷ The four C-terminal residues of the 2-fold-related cro monomers, which are disordered in the crystals, and also, presumably, in solution, could also interact with the DNA (Figs 5, 6, 7). Such interactions could include Lys 62 and Lys 63. Presumably these C-terminal arms, which extend in solution like 'feelers', occupy well defined positions when complexed with the DNA. The finer details of the interactions between the cro repressor and its specific operator sequences are now being explored by more careful model building, and will be described elsewhere.

The proposed binding of cro to DNA explains many of the known properties of the complex. A tetramer or a dimer of cro bound to DNA extends over exactly 17 base pairs, the length of one operator (Figs 3, 5). The mode of interaction between protein and DNA is consistent with the approximate 2-fold symmetry of the operator sequence, and the symmetry of the protein. Also, the proposed binding explains, within the present level of detail, the chemical protection experiments (Figs 3, 6). Furthermore, there is a space of ~20 Å between cro molecules bound to adjacent operators, consistent with non-cooperativity of binding.²⁸

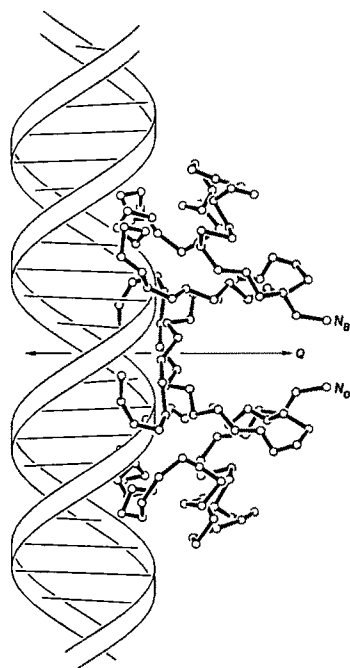


Fig. 5 Presumed interaction of cro repressor with DNA. Two monomers of cro (O and B), related by 2-fold symmetry axis Q, interact with the DNA so that axis Q coincides with the 2-fold symmetry axis of the DNA. The DNA has been rotated 90° relative to that in Fig. 3 so that its 2-fold symmetry axis is in the plane of the paper. The respective amino termini of the two cro molecules are labelled N_O and N_B . A pair of 2-fold-related α -helices occupy successive major grooves of the DNA and, closer to the symmetry axis, two extended polypeptide strands run parallel to the backbone of the DNA.

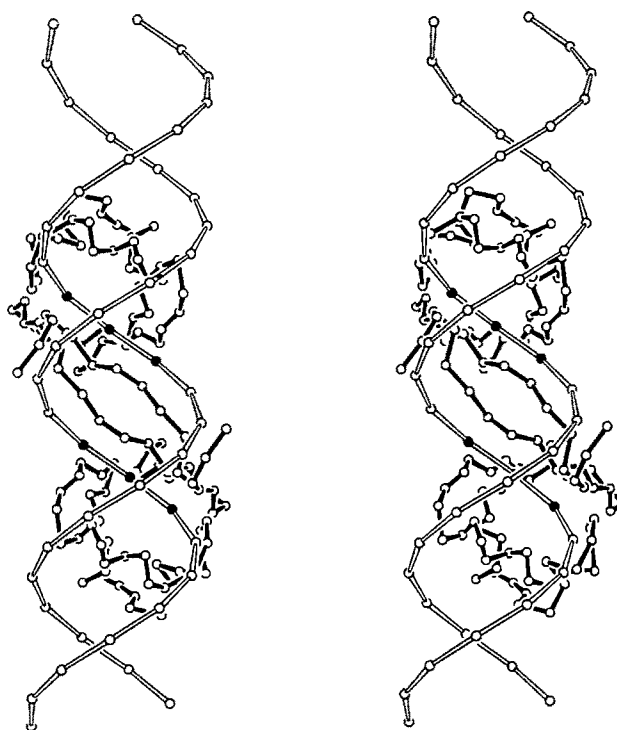


Fig. 6 Stereo view of the presumed interaction between two molecules of cro repressor and DNA, viewed along their common 2-fold symmetry axis. The open circles along the DNA backbone indicate the positions of the phosphates, those phosphates whose ethylation hinders cro binding (Fig. 3) are drawn solid.

Conclusions

This is the first structure determination of a specific DNA repressor, and the results obtained here may be of general relevance for interactions between proteins and DNA.

A principal feature of the proposed binding of cro to DNA is the use of symmetry. The 2-fold symmetry axis of the repressor coincides with the 2-fold symmetry axis of the DNA operator. This type of interaction is expected to be of general significance for proteins which recognize symmetric, sequence-specific regions of the DNA. Departures from exact symmetry could be important in 'fine tuning' of the strength of the DNA-protein complex. Our results indicate that the symmetric sequences often used for specific DNA-protein interactions are a consequence of the oligomeric structure of the regulatory proteins rather than the source of a special type of DNA structure.

Also, the α -helical conformation of proteins may presumably be used in both specific and nonspecific DNA binding. Sung and Dixon²⁹ were the first to propose that the histone H4 N-terminus adopts a helical conformation when interacting with chromatin nucleic acid. A related model has been proposed for *lac* repressor³⁰. Also, Warrant and Kim³¹ have presented direct evidence showing that protamine changes to an α -helical conformation on binding to a double-helical portion of crystalline tRNA. The cro repressor structure supports the notion that the binding of α -helices in the major groove of DNA may be a principal determinant in nonspecific protein binding to DNA, and is also of importance in specific recognition.

The structure of cro, and its complementarity to the B-form of DNA, suggest that the DNA retains the right-handed B-conformation when the repressor is bound. Of course, slight deformations of the DNA or protein could well occur on binding, and may be suggested by detailed model-building experiments, but the present results indicate that large adjustments in the DNA structure are not required to accommodate the protein. In contrast, McKay and Steitz³² have recently proposed that the cyclic AMP receptor protein of *Escherichia coli*, which promotes the action of RNA polymerase, binds to left-handed DNA. In the case of the cro repressor, the protein is located

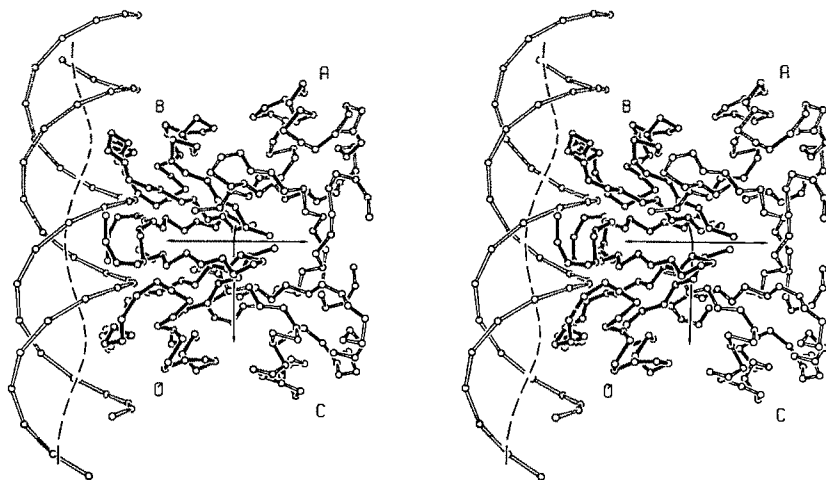


Fig. 7 Stereo view of the presumed binding of a tetramer of cro to DNA. Monomers O and B, seen interacting with the DNA, are drawn with solid bonds, whereas monomers A and C which form a second potential DNA binding region on the opposite side of the tetramer, are drawn with open bonds. The broken line follows the 'bottom' of the major groove of the DNA. The DNA and protein have been rotated 30° about the horizontal axis so that the direction of view is essentially along the major groove of the DNA.

against one side of the DNA, with most of the specific recognition occurring between side chains of the protein and the parts of the bases that are exposed in the major groove of the DNA. There is no apparent need to expose the bases further to facilitate recognition.

We thank Drs Howard Schachman and Harrison Echols for their help in instigating this project. We also thank our

colleagues, in particular D. Bacon, R. G. Fisher, I. Herskowitz, M. F. Schmid, G. R. Smith, L. F. Ten Eyck and L. H. Weaver, for assistance and advice. This work was supported in part by grants from the NSF (PCM-8014311), the NIH (GM 20066, GM 28138), the MRC of Canada and the M. J. Murdock Charitable Trust. D. H. O. was supported by a Damon Runyan-Walter Winchell Cancer Fund Postdoctoral Fellowship (DRG-303-F).

Received 10 December 1980, accepted 18 February 1981

- 1 Echols, H. *A Rev Genet* **6**, 157-190 (1972)
- 2 Herskowitz, I. *A Rev Genet* **7**, 289-323 (1973)
- 3 Ptashne, M. *et al Science* **194**, 156-161 (1976)
- 4 Ptashne, M. *et al Cell* **19**, 1-11 (1980)
- 5 Johnson, A., Meyer, B. J. & Ptashne, M. *Proc natn Acad Sci USA* **75**, 1783-1787 (1978)
- 6 Takeda, Y. *J molec. Biol* **127**, 177-191 (1979)
- 7 Folkmanis, A., Takeda, Y., Smith, J., Gussin, G. & Echols, H. *Proc natn Acad Sci USA* **73**, 2249-2253 (1976)
- 8 Takeda, Y., Folkmanis, A. & Echols, H. *J biol Chem* **252**, 6177-6183 (1977)
- 9 Hsiang, M. W., Cole, R. D., Takeda, Y. & Echols, H. *Nature* **270**, 275-277 (1977)
- 10 Roberts, T. M., Shimatake, H., Brady, C. & Rosenburg, M. *Nature* **270**, 274-275 (1977)
- 11 Green, D. W., Ingram, V. M. & Perutz, M. F. *Proc R Soc A* **225**, 287-307 (1954)
- 12 Matthews, B. W. in *The Proteins* Vol. 3 3rd edn (eds Neurath, H. & Hill, R. L.) 403-590 (Academic, New York 1977)
- 13 Blundell, T. L. & Johnson, L. N. *Protein Crystallography* (Academic, London, 1976)
- 14 Anderson, W. F., Takeda, Y., Echols, H. & Matthews, B. W. *J molec Biol* **130**, 507-510 (1979)
- 15 Anderson, W. F., Takeda, Y., Ohlendorf, D. H. & Matthews, B. W. *Proc 7th Katzir-Kachalsky Conf on Structural Aspects of Recognition and Assembly in Biological Macromolecules* March 1980 (in the press)
- 16 Schmid, M. F. *et al Acta crystallogr* (in the press)
- 17 Molnar, C. E., Barry, C. D. & Rosenberger, F. U. *Tech Memo No 229* (Computer Systems Laboratory, Washington University, St. Louis, 1976)
- 18 Blow, D. M. & Crick, F. H. C. *Acta crystallogr* **12**, 794-802 (1959)
- 19 Brice, G. *Acta crystallogr* **A30**, 395-405 (1974)
- 20 Richards, F. M. *J molec Biol* **37**, 225-230 (1968)
- 21 Colman, P. M., Jansonius, J. N. & Matthews, B. W. *J molec Biol* **70**, 701-724 (1972)
- 22 Maniatis, T. *et al Cell* **5**, 109-113 (1975)
- 23 Flashman, S. M. *Molec gen Genet* **166**, 61-73 (1978)
- 24 Meyer, B. J., Maurer, R. & Ptashne, M. *J molec Biol* **139**, 163-194 (1980)
- 25 Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737 (1953)
- 26 Arnott, S. & Hukins, D. W. L. *Biochem biophys Res Commun* **47**, 1504-1509 (1972)
- 27 Church, G. M., Sussman, J. L. & Kim, S.-H. *Proc natn Acad Sci USA* **74**, 1458-1462 (1977)
- 28 Johnson, A., Meyer, B. J. & Ptashne, M. *Proc natn Acad Sci USA* **76**, 5061-5065 (1979)
- 29 Sung, M. T. & Dixon, G. H. *Proc natn Acad Sci USA* **67**, 1616-1623 (1970)
- 30 Adler, K. *et al Nature* **237**, 322-327 (1972)
- 31 Warrant, R. W. & Kim, S. H. *Nature* **271**, 130-135 (1978)
- 32 McKay, D. & Steitz, T. A. *Nature* (this issue)
- 33 Herskowitz, I. & Hagen, D. A. *Rev Genet* **14**, 399-445 (1980)
- 34 Echols, H. in *The Molecular Genetics of Development* (eds Loomis, W. & Leighton, T.) 1-16 (Academic, New York, 1980)
- 35 Johnson, A. thesis, Harvard Univ (1980), and manuscript in preparation

LETTERS

Solar quiescent prominences at 10.7 GHz

R. N. Bracewell

Electrical Engineering Department, Stanford University, Stanford, California 94305, USA

W. Graf

SRI International, Menlo Park, California 94025, USA

Filaments on the Sun, which are dark linear features in the light of the Fraunhofer H α line, are a mysterious feature of the solar surface. They are blade-like extensions of the chromosphere projecting to a height of 40,000 km into the corona, and may extend for up to 200,000 km along the surface. The fact that they may persist for a year or so, disappear abruptly and then reappear in the same positions shows that they are superficial indicators of deeper phenomena. There have been several reports¹⁻⁴ of reduced brightness in the radio spectrum associated with the H α filaments, leading to the widely accepted view that filaments are associated with broad dark radio features. This picture is, however, contradicted by our observations at 10.7 GHz of one solar filament during the solar eclipse of 12 October 1977.

Although the filaments are very thin, subtending about 8 geocentric arc s when seen vertically, nevertheless, "regions of reduced radio brightness situated above filaments" were detected¹ in 1959 and explained in terms of "relatively cold prominences which are optically thick to 8-mm radiation". Later attempts to discern filaments at radio wavelengths succeeded due to use of the 11 m mm-wave telescope at Kitt Peak and the 42 m telescope at Green Bank²⁻⁴. In all cases temperature depressions in the neighbourhood of filaments were reported. Khangil'din's¹ interpretation was absorption in the filamentary material, Drago and Felli⁴ suggested that the temperature defect arose from lack of coronal emission from the volume occupied by the filament. The width of the depression was confirmed at ~2 arc min by Straka *et al*⁵ who proposed a model where the filament would exhibit a greater brightness temperature because the enhanced density within the filament would raise the height of the emission region. The general depression previously reported would then be due to the low-density 'cavity' surrounding the filament. Observations⁶ with the 100-m telescope at Effelsberg again confirmed the temperature depression. The depression was explained in terms of 'radio filament' for the 'temperature depressions' or 'absorption features' even though the width was perhaps five times larger than that of the dark optical filament. This interpretation in terms of absorption disagreed with Straka *et al*'s model⁵ of the cavity concept because Kundu *et al*⁶ believed that the cavity model

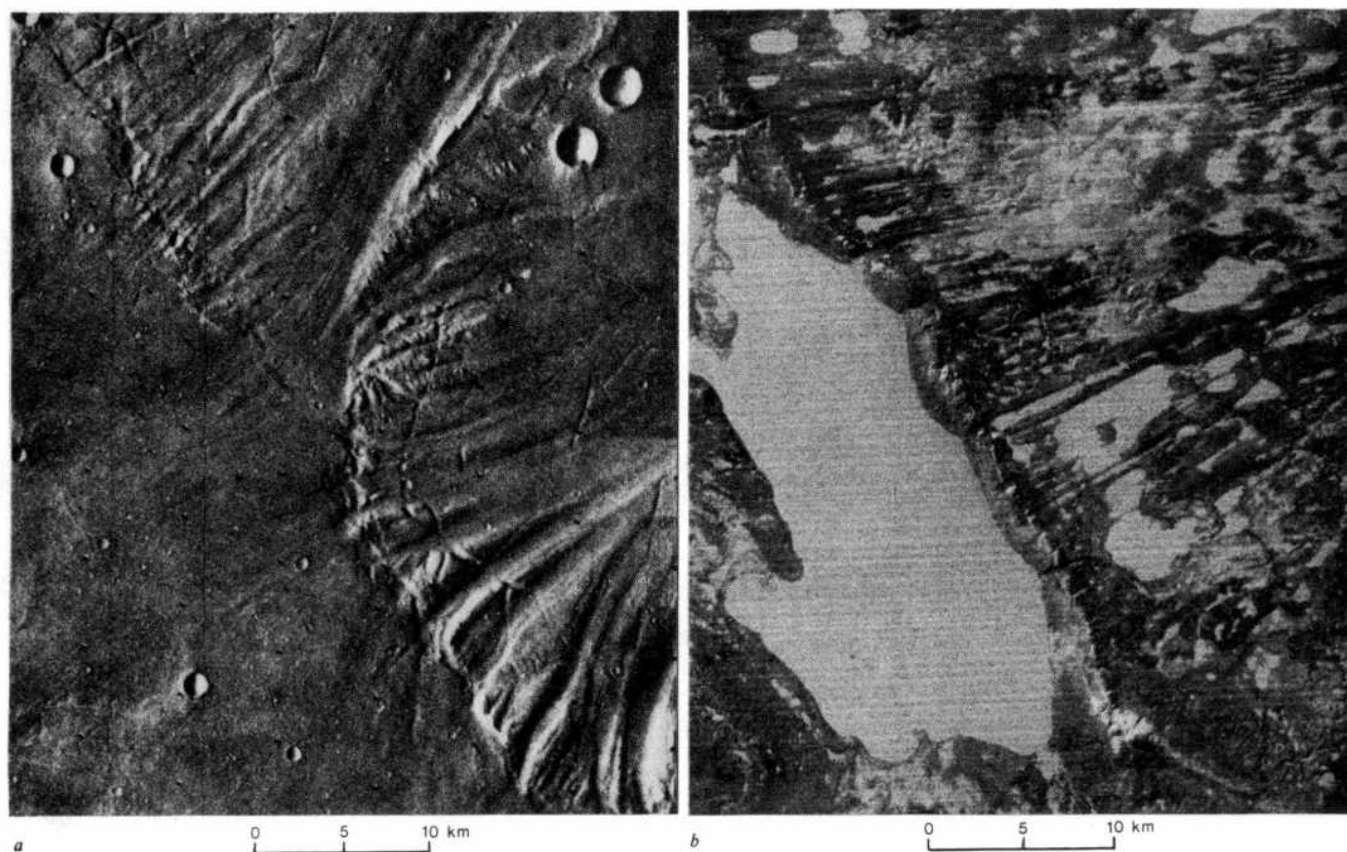


Fig. 2 Grooved terrain on Mars and Earth. Cross ridges in centre of images apparently dammed and ponded carving material, which eventually breached and deeply scoured the terrain at right. Flow was towards the right. *a*, Breached wrinkle ridge in Chryse Planitia and channel scour marks. North is towards the top. (Viking 46A65). *b*, Scarp near Lac Belot, Northwest Territories, Canada, and glacial grooves. North is towards the bottom. (Landsat 1594-19302-7).

measured widths of as much as 180 km across the grooved terrain of Laurentide ice sheets in the Northwest Territories of Canada. The comparable width of grooved terrain in Kasei Vallis and Canada suggests that some of the ice on Mars may have been in sheets rather than valley glaciers.

Scarp heights of the martian channels also compare favourably with terrestrial glacial scarps. Scarps flanking Kasei Vallis lie near 2,000 m, and those at the mouth of Tiu and Ares Valles near 600 m. Grooves on the level ground above and below some of the scarps indicate that the ice may locally have inundated the scarps and may have been thicker than 2,000 m. Such heights are not excessive for terrestrial glacial terrain; scarp heights on Lambert Glacier in Antarctica are as much as 3,000 m, and the glacier presently is thicker than 2,000 m (ref. 11). Ice streams into the Ronne Shelf near the Ellsworth Mountains are 1,000–2,000 m thick¹² (Fig. 3). Pleistocene ice streams in Iceland are deduced to have been 1,000 m thick (based on the height of tablemountains)¹³, and the present-day Antarctic and the former Laurentide shields are 3,000–4,000 m thick¹⁴.

To determine whether ice sheets would have moved in martian conditions, it is important to obtain information on ice-surface gradients. Because no ice sheets are presently in the areas of interest, only the gradients of the martian channel beds can be obtained. For this discussion, however, ice-surface gradients can be assumed to be parallel. Regional gradients are very low and range from $<0.1^\circ$ to 0.25° ; local segments are steeper. Low ground gradients are no deterrent to glacial movement on Earth, however. For instance, the Yenta River in Alaska has a gradient of 0.06° , and the Wilkes subglacial basin in Antarctica has a gradient of $\sim 0^\circ$ for nearly 2,000 km. Thus, ground gradients apparently are not a critical determinant for glacial motion; indeed, glaciers are known to move uphill under certain circumstances. Terrestrial ice-sheet surface gradients range from 0.10° to 0.25° , from the Antarctic highs to outlet

glaciers on the coast, and thus are comparable to the gradients of martian channel beds.

It remains to be considered whether or not ice could have moved through martian outflow channels, given present and past martian environmental conditions. Sharp¹⁵ argued against the likelihood of glacial motion on Mars; however, he only considered the cold polar areas and not the warmer equatorial belt, where most outflow channels occur. Existing data on ice-deformation processes have been collected and displayed by Shoji and Higashi¹⁶ in a plot of shear stress versus temperature. Taking nominal values of surface slope (0.1 – 0.25°) and ice thickness (1,500–3,000 m), basal shear stresses (τ) may be calculated from the relation

$$\tau = \rho g y \sin \alpha \quad (1)$$

where ρ , g , and y are the ice density, gravitational acceleration, and depth, respectively, and α is the surface slope. When these shear stresses are plotted against martian equatorial temperatures (0° to -60°C), the conventional terrestrial processes characteristic of ice-sheet deformation—namely, dislocation creep and lattice-diffusion creep—also seem to apply to ice deformation on Mars. In addition, the shear stress–temperature field that is defined for Mars includes all the measured shear stress–temperature data obtained at various depths in drill holes at Byrd Station, Antarctica, and Camp Century, Greenland. These results strikingly illustrate the similarity in conditions on the two planets, notwithstanding the generally lower temperatures now prevailing on Mars.

Nominal strain rates, given in the plot of shear stress versus temperature by Shoji and Higashi¹⁶, vary by many orders of magnitude and diminish with decreasing temperatures and shear stress. These data suggest that ice, in present-day martian conditions, would have moved only very slowly. Several other factors, however, could increase the strain rate. Preferred crystal

Fig. 3 Streamlined island on Mars and Earth. Flow diverging around obstacles formed by islands on Mars and ice rises on Earth. Martian islands tend to be smaller than ice rises. Flow is towards the bottom. *a*, Kasei Vallis. North is towards the right. (Viking 519A29). *b*, Ice rises in Ronne ice shelf, Antarctica. Ellsworth Mountains on left. Radio echo sounding suggests subglacial breaching of ice rises by fjordlike troughs, which would further enhance their similarity to martian islands, north is towards the top right. (Satellite image map, Ellsworth Mountains, Antarctica, US Geological Survey, original scale 1:500,000).



orientation and occluded gas bubbles in polycrystalline ice sheets act to increase strain rates at given temperatures and shear stresses by as much as 100-fold^{17,18}. Strain rates can also be increased once flow by dislocation creep is initiated by local steeper slopes and the ice mass undergoes frictional heating. The energy thus released internally by an ice sheet flowing at a velocity of 20 m yr^{-1} is roughly equal to the normal geothermal flux. An ice velocity of 20 m yr^{-1} is not extraordinary and, on

Earth or Mars, could be considerably larger if ice-surface slopes increased to several degrees. Artesian water under the ice could also have contributed to lubrication of the bases of martian ice masses and thus accelerated their motion and increased their erosive power. The pronounced grooving on the channel floors indeed suggests that the ice moved over the ground and eroded powerfully. A warmer climate in the past, of course, would have favoured faster glacial motion.

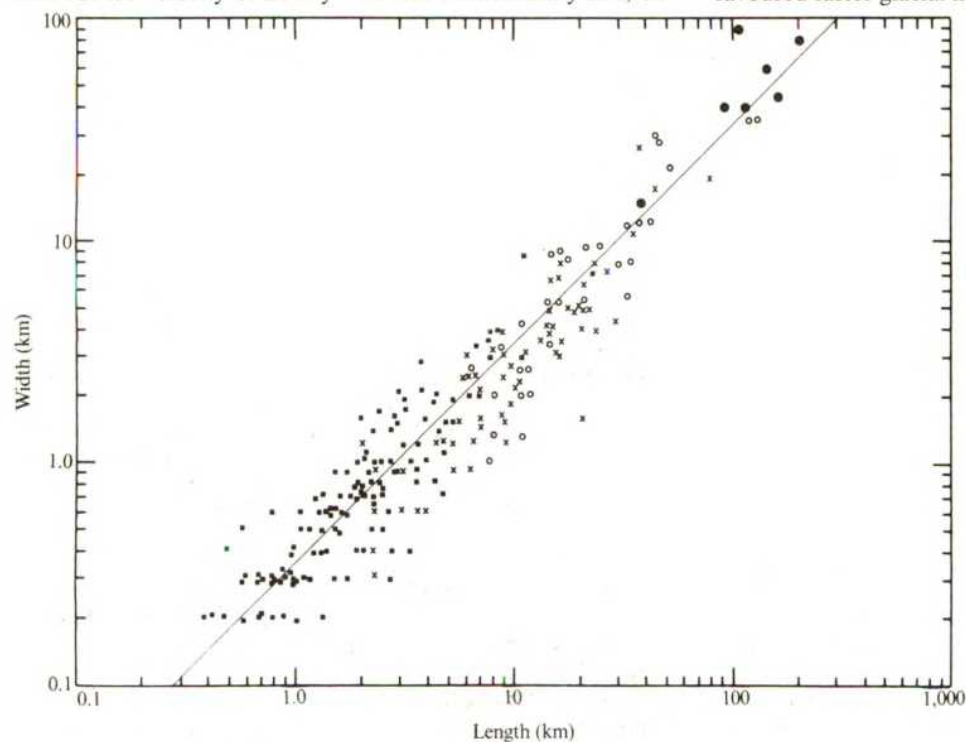


Fig. 4 Dimensions of streamlined islands. Plotted scabland flood islands (■) and martian islands are from ref. 7; ×, Maja Vallis; ○, Kasei Vallis. Plotting of Antarctic ice-rise dimensions (●) on same trendline as scabland and Mars values indicates that sculpturing by ice is as likely as catastrophic flooding for causing streamlining.

Ice may have accumulated in the source area of the martian channels from atmospheric precipitation, wind-drifted frost and dust deposits, and from the ground as springs. All these processes no doubt contributed, but springs must have been dominant because the chaotic terrain, the source area of the channels, is highly disturbed, apparently owing to the removal of large quantities of material. The groundwater may have been under artesian pressure as proposed by Carr¹⁹, and created ice domes that then flowed to adjacent depressions to carve the channels. Alternatively, the water could have formed ice covers over incipient rivers that required only normal discharges²⁰. The river ice may have consolidated into ice drives⁷ and, eventually, ice streams. The internal deformation and flow of such ice streams may then have imparted to the martian channels their distinctive and spectacular morphology.

We thank H. Masursky, M. Grolier and V. Baker for helpful reviews (although their point of view is not necessarily that of the authors). This work was supported by NASA contract W14,575 and grant 150-6330A.

Received 22 December 1980; accepted 8 March 1981.

1. Sharp, R. P. & Malin, M. C. *Bull. geol. Soc. Am.* **86**, 593–609 (1975).
2. Pieri, D. C. *Science* **210**, 895–897 (1980).
3. Baker, V. R. & Milton, D. J. *Icarus* **23**, 27–41 (1974).
4. McCauley, J. F. *et al. Icarus* **17**, 289–327 (1972).
5. Masursky, H., Boyce, J. M., Dial, A. L., Schaber, G. G. & Strobbe, M. E. *J. geophys. Res.* **82**, 4016–4038 (1977).
6. Baker, V. R. *Science* **202**, 1249–1256 (1978).
7. Baker, V. R. *J. geophys. Res.* **84**, 7985–7993 (1979).
8. Nummedal, D. in *NASA Tech. Mem.* 79729, 257–259 (1978); **81776**, 289–291 (1980).
9. Thompson, D. E. *Geophys. Res. Lett.* **6**, 735–738 (1979).
10. Baker, V. R. & Kocheil, R. C. *Proc. 9th Lunar planet. Sci. Conf.* 3193–3203 (1978).
11. Morgan, V. I. & Budd, W. F. *J. Glaciol.* **15**, 103–111 (1975).
12. Swinbank, C. *Phil. Trans. Soc. B* **279**, 161–183 (1977).
13. Preusser, H. *Landscapes of Iceland: Types and Regions* (Junk, The Hague, 1976).
14. Sugden, D. E. *Arctic Alp. Res.* **9**, 21–47 (1977).
15. Sharp, R. P. *J. Glaciol.* **13**, 173–185 (1974).
16. Shoji, H. & Higashi, A. *J. Glaciol.* **21**, 419–427 (1978).
17. Garfield, D. E. & Ueda, H. T. *J. Glaciol.* **17**, 29–34 (1976).
18. Whillans, I. M. *J. Glaciol.* **18**, 359–372 (1977).
19. Carr, M. H. *J. geophys. Res.* **84**, 2995–3007 (1979).
20. Wallace, D. & Sagan, C. *Icarus* **39**, 385–400 (1979).

Decade fluctuations in geomagnetic westward drift and Earth rotation

Jean-Louis Le Mouél*, Theodore R. Madden†, Joël Ducruix* & Vincent Courtillot‡*

* Institut de Physique du Globe, Université Paris VI, 75230 Paris Cedex 05, France

† Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

‡ UER des Sciences Physiques de la Terre, Université Paris VII 75230 Paris Cedex 05, France

Numerous studies of the variations of the Earth's rotation have been made since 1900 (summarized in refs 1, 2). It seems well established that variations in the rotation with periods equal to or less than 2 yr are due to variations in the zonal circulation of the atmosphere^{2–4}. However, there is no general agreement about the cause of the longer-period variations of the Earth's rotation, sometimes known as the 'decade variations', although the most favoured explanation is electromagnetic core–mantle coupling (see ref. 5). Several studies^{6–8} of the correlation between Earth rotation and magnetic field variations have obtained marginally positive results. We have been able to clarify such a correlation over the past 120 yr by studying those magnetic field indicators that are the most sensitive to variations in westward drift of the outer regions of the core. We suggest here that the correlation can be extended back ~300 yr. Simple models of core–mantle coupling, based on Bullard's ideas, give reasonable quantitative agreement with the observations.

In 1951, Stoiko⁹ compared the variations of the Earth's magnetic field intensity observed at Paris with the variations of

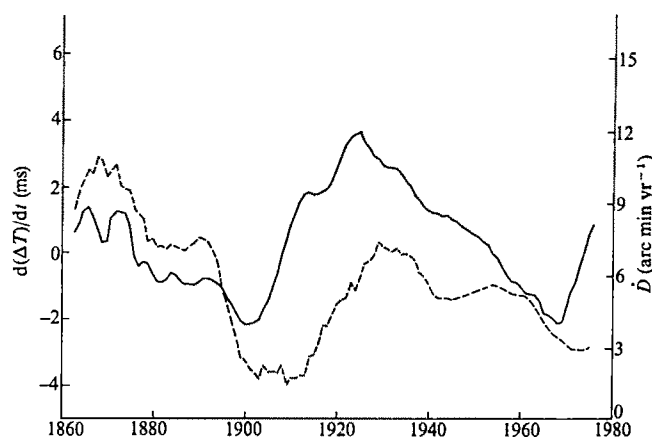


Fig. 1 Plot of the secular variation of declination $\dot{D}$ (solid curve and right hand scale) and of the excess length of day $d(AT)/dt$ (dashed curve and left-hand scale) for the period 1861–1978. ΔT data are from refs. 18, 19. The relative changes in the Earth rotation rate $\Delta\Omega/\Omega$ are obtained from the $d(AT)/dt$ values through multiplication by -1.157×10^{-8} .

the length of day for the period 1883–1948 and claimed a good correlation between the two. In 1953 Vestine⁶ proposed a correlation between the westward drift of the eccentric dipole (the magnetic centre C) and the variations of the Earth's rotation rate Ω for the period 1885–1945, and gave a qualitative interpretation of this result using Bullard's model. Vestine and Kahle¹⁰ extended the correlation up to 1965. A puzzling feature of this correlation is that the magnetic variations seem to lag behind the Ω variations. The variations of the longitude of C for the period 1945–75 have been found¹¹ to have a poor correlation with Ω . More recently, it has been shown^{12,13} that, starting in 1969, there was a rapid acceleration of the secular variation of the east (Y) component of the magnetic field across large regions of the Northern Hemisphere and that a similar event occurred around 1900 (although it is more difficult to ascertain its global nature because of the lack of magnetic observatories at that time). Minima in Ω were also noted at about the same times. It is this correlation between the Y variations and Ω which we consider and extend here.

Motions of the top layers of the core are responsible for most magnetic variations seen at the Earth's surface on the decade time scale; the westward drift of magnetic features which has long been noted^{14,15} can be taken as evidence for a westward drift of the upper core. Hide¹⁶ has pointed out that such motions could be the result of slow MHD–Rossby waves, but the simultaneity of the changes in these motions seen around the world makes this possibility unlikely.

To study the westward drift of magnetic features we have to distinguish between the drift effect and other causes of magnetic variations¹⁷. The magnetic secular variation can be written as:

$$\frac{\delta B}{\delta t} = \frac{\partial B}{\partial t} + U_{\phi} \frac{\partial B}{\partial \phi}$$

where $\delta B/\delta t$ is the magnetic variation seen at magnetic observatories; the $\partial B/\partial t$ term accounts for the effects of non uniform and north–south motions (as well as the less important effects of magnetic diffusion), U_{ϕ} is the magnitude of the westward drift as seen at the Earth's surface and $\partial B/\partial \phi$ is the longitudinal gradient of the magnetic field as seen at the surface.

Declination D (or the closely associated east component Y) is especially suited to studies of zonal motions because (1) it has weak gradients in the north–south directions compared with the other components and (2) it was more accurately and frequently determined in earlier times.

To study the detailed time history of the magnetic secular variations data from continuously recording magnetic observatories are needed. Most of the early magnetic observatories are located in Europe and it is fortunate that here the $\partial Y/\partial \phi$ term is large and smoothly varying. Thus $\delta Y/\delta t$ is dominated, at

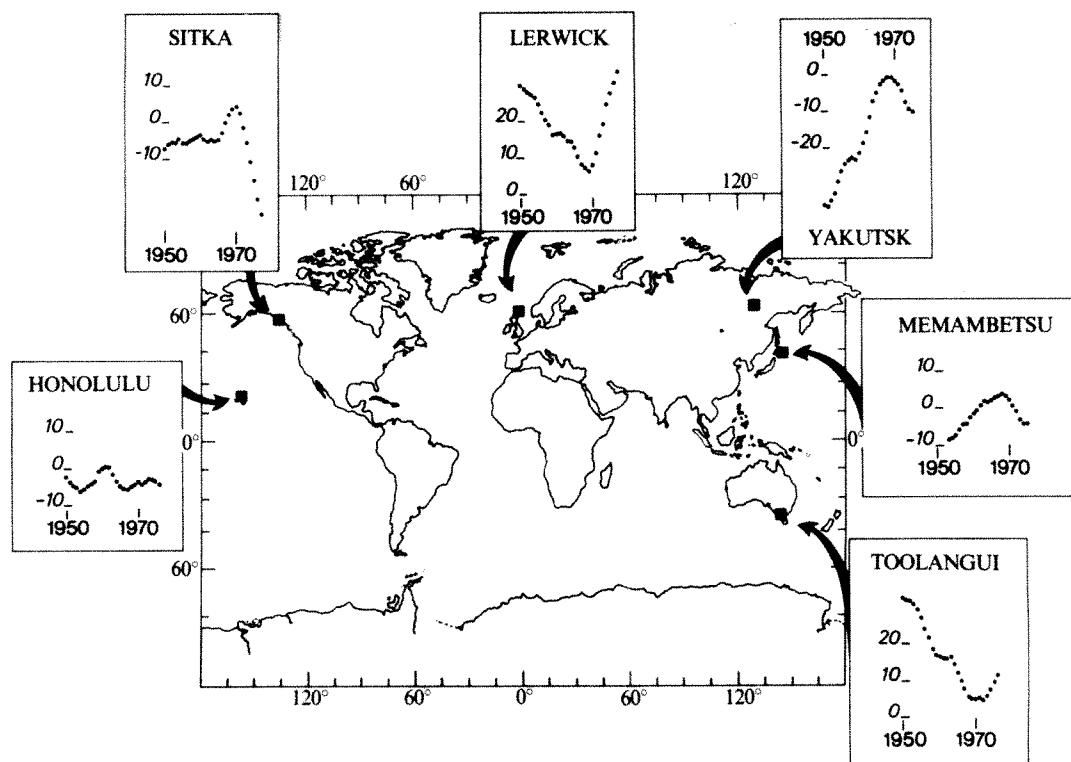


Fig. 2 Selected $\delta Y/\delta t$ variations (in nT yr^{-1}) around the period of apparent westward drift minimum of 1969.

least for the past century, by the term $U_\phi(\partial Y/\partial \phi)$, and $\dot{Y}$ (or $\dot{D}$) should be proportional to U_ϕ . For Paris $\partial Y/\partial \phi$ is $\sim 150 \text{ nT per deg}$ and $\partial D/\partial \phi$ is $\sim 0.4\text{--}0.5$. If core-mantle coupling is the cause of the decade variations in Earth rotation one should expect a good correlation between $\dot{Y}$ (or $\dot{D}$) as seen in Europe and Ω .

Figure 1 displays $\dot{D}$ and $\Delta\Omega/\Omega$ for the period 1861–1978. The D values are from a compilation of available data from the Paris area. These data can be considered as representative of the western European area, as has been shown for the recent observatory data¹³. $\dot{D}$ has been averaged over 5 yr. The Ω data for 1861–1978 are from refs 18, 19.

The data of Fig. 1 show a good correlation between $\dot{D}$ and Ω , significantly better than in any previous study (see Fig. 4 of ref. 10 or Fig. 4 of ref. 8). The magnetic variations lead, rather than lag⁷, the Ω variations by $\sim 10 \text{ yr}$. (The coherency peaks fairly sharply at about 0.8 for this value of the lag. A more sophisticated correlation analysis is not warranted for the kind of data shown in Fig. 1 and the eye is still the best judge.) These variations are quite pronounced. From $\dot{D}$ one can infer that the westward drift decreased from ~ 0.3 to 0.1 deg yr^{-1} between 1870 and 1900, increased again from ~ 0.1 to 0.4 deg yr^{-1} between 1900 and 1925 and decreased again to $\sim 0.1 \text{ deg yr}^{-1}$ in 1969 (see also ref. 11). Since then a rapid acceleration of the drift has been seen, which is quite similar to that observed just after 1900.

The most pronounced features of secular variation are the discontinuities in the slope of $\dot{Y}$ seen in 1900 and 1969, and somewhat less clearly in 1925, which we interpret as sudden changes in the acceleration of the westward drift, $\dot{U}_\phi$. The number and distribution of magnetic observatories in 1900 were too limited to ascertain the global character of the drift acceleration changes, but by 1969 the situation was greatly improved. Fig. 2 shows selected results from a preliminary study, still in progress, which shows that in most areas where we expect to see the effects of rapid changes in the slope of U_ϕ they are indeed observed. The examples in Fig. 2 come both from regions where $\partial Y/\partial \phi$ is large and smoothly varying (for example, Europe^{12,13}) and from regions where $\partial Y/\partial \phi$ is close to zero (for example, Honolulu): in the former a clear extremum of

$\partial Y/\partial \phi$ is observed, in the latter there is none. However, in some cases (for example, South America) discrepancies are found which may be due to the evolution of more localized magnetic features. In any case, the magnetic data are not incompatible with a global zonal circulation pattern existing in the upper core which underwent a rapid change of acceleration in 1969^{13,20}.

We can only attempt here to identify some of the parameters that are relevant in a quantitative study of core-mantle coupling and estimate their magnitudes. We follow Bullard's model of a convecting core which causes differential rotation between the outer regions and the inner regions. This differential rotation creates a toroidal magnetic field in the core which penetrates into the lower mantle¹⁵. Thus, the mantle, upper and inner core are all electromagnetically coupled, and, if the convection regime changes, the core and mantle will all undergo changes in their angular velocities, the net angular momentum remaining constant. When the upper core region is thin, the mantle follows the inner core and will accelerate when the upper core decelerates (see ref. 20, with lower mantle conductivity estimates from ref. 21): this is what is observed. The thickness of the upper core zone can be estimated by assuming that the mantle and lower core are rotating together. Using the law of conservation of momentum and the observed proportionality rate between $\delta\Omega$ and δU_ϕ (with $\delta\Omega \sim 3 \times 10^{-12} \text{ rad s}^{-1}$ and $\delta U_\phi \sim 0.2 \text{ deg yr}^{-1}$) a thickness of 100 km is found for this upper region. The time lag of the changes in Ω relative to U_ϕ is a measure of the magnitude of the coupling between the mantle and the inner core, which in turn depends on the magnitude of the toroidal field in the core and the conductivity of the lower mantle. The observed time lag can be accounted for by the above model (Le Mouél and Courtillot, in preparation).

We have also compared data for the period 1700–1870. Despite low reliability, these older data do not display the dramatic changes that have occurred since 1870. These observations seem to imply that the period 1700–1870 was a more stable regime for core motions than the following century. The similarity of the magnetic situation in 1900 and 1969 leads us to predict that the Earth rotation is presently at the beginning of a period of acceleration.

Received 1 December 1980; accepted 20 February 1981.

1. Munk, W. H. & McDonald, G. J. F. *The Rotation of the Earth* (Cambridge University Press, London 1960).
2. Lambeck, K. *The Earth's Variable Rotation. Geophysical Causes and Consequences* (Cambridge University Press, London, 1980).
3. Cazenave, A. thesis, Toulouse, Univ. (1975).
4. Lambeck, K. & Cazenave, A. *Phil. Trans. R. Soc. A* **284**, 495–506 (1977).
5. Runcorn, S. K., in *Earthquake Displacement Fields and the Rotation of the Earth* (eds. Mansinha, L. et al.) 181–187 (Reidel, Dordrecht, 1970).
6. Vestine, E. H. *J. geophys. Res.* **58**, 127–145 (1953).
7. Kahle, A. B., Ball, R. H. & Cain, J. C. *Nature* **223**, 165 (1969).
8. Yukutake, T. *J. Geomagn. Geoelectr.* **25**, 195–212 (1973).
9. Stoyko, N. C. *r. hebdom. Séanc. Acad. Sci., Paris* **205**, 1577–1579 (1937); **223**, 80–82 (1951); **224**, 1798–1800 (1952).
10. Vestine, E. H. & Kahle, A. B. *Geophys. J. R. astr. Soc.* **15**, 29–37 (1968).
11. Harwood, J. M. & Malin, S. R. C. *Nature* **259**, 469–471 (1976).
12. Courtillot, V., Ducruix, J. & Le Mouél, J. L. *C. r. hebdom. Séanc. Acad. Sci. Paris D* **287**, 1095–1098 (1978); **B289**, 173–175 (1979).
13. Ducruix, J., Courtillot, V. & Le Mouél, J. L. *Geophys. J. R. astr. Soc.* **61**, 73–94 (1980).
14. Halley, E. *Phil. Trans. R. Soc.* **12**, 208–221 (1863).
15. Bullard, E. C., Freedman, C., Gellman, H. & Nixon, J. *Phil. Trans. R. Soc. A* **243**, 67–92 (1950).
16. Hide, R. *Phil. Trans. R. Soc. A* **259**, 615–650 (1966).
17. Yukutake, T. *Bull. Earthq. Res. Inst. Tokyo* **40**, 1–64 (1962).
18. Morrison, L. V. *Geophys. J. R. astr. Soc.* **58**, 349–360 (1979).
19. Guinot, B. *La Terre, les Eaux, l'Atmosphère* Ch. 2 (Rotation de la Terre, Bureau des Longitudes, Paris, 1977).
20. Le Mouél, J. L. & Courtillot, V. *Phys. Earth planet. Inter.* (in the press).
21. Achache, J., Courtillot, V., Ducruix, J. & Le Mouél, J. L. *Phys. Earth planet. Inter.* **23**, 72–75 (1980); *Geophys. J. R. astr. Soc.* (in the press).

Phonon scattering at grain boundaries in heavily doped fine-grained silicon-germanium alloys

D. M. Rowe, V. S. Shukla & N. Savvides*

Department of Physics, Electronics and Electrical Engineering, UWIST, Cardiff CF1 3NU, UK

Thermocouples made from heavily doped n- and p-type silicon-germanium alloys are used in the nuclear powered thermoelectric generators which provides onboard power to the American Voyager spacecraft. The isotopic heat source is expensive and any improvement in the generators' heat energy to electrical energy conversion efficiency would save both money and weight. The possibility of reducing the thermal conductivity of thermoelectric semiconductor alloys by compacting fine grained material has been considered and various estimates have been made of the reduction in thermal conductivity with grain size. The electrical properties of hot pressed compacts of heavily doped silicon germanium alloys which possess density values close to that of single crystal are unaffected by grain boundary scattering effects. Consequently, this method of preparation holds out a possibility of improving the 'figure of merit' of the thermocouple material and hence the conversion efficiency of the generator. Here the first measurements of the thermal diffusivity of fine grained hot pressed compacts of heavily doped n-type silicon germanium alloy are reported. The compacts investigated possessed grain sizes (L) in the range $10 < L < 25 \mu\text{m}$, $5 < L < 10 \mu\text{m}$ and $< 5 \mu\text{m}$. The results confirm that the lattice thermal conductivity decreases with a reduction in grain size. In the compact with $L < 5 \mu\text{m}$, this reduction is $\sim 28\%$ compared with 'single crystal' material. Within experimental error ($< 5\%$) the Seebeck coefficient and electrical resistivity do not change with grain size. This implies that the conversion efficiency of a thermoelectric generator employing silicon-germanium thermocouples would be significantly improved through the use of fine-grained thermocouple material.

The conversion efficiency of a semiconductor thermocouple depends on the temperature difference over which the device operates, its average temperature and the so-called figure of merit Z for the thermocouple material. $Z = S^2\sigma/K$ where K is the thermal conductivity, S the Seebeck coefficient and σ the

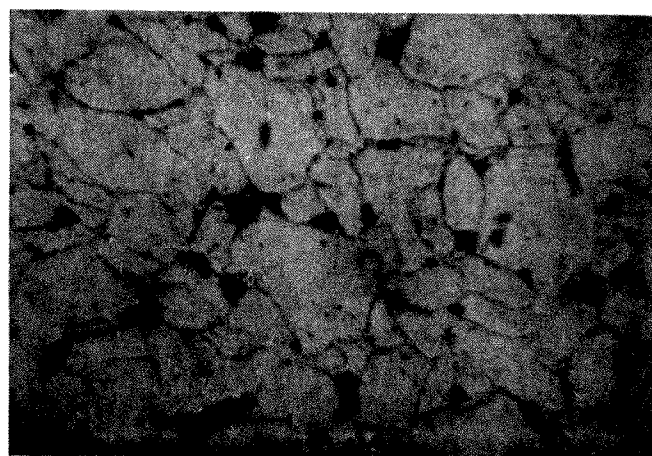


Fig. 1 Microphotograph of a typical polished surface of a compact after etching with CP4. Grain size is in the range $10 < L < 25 \mu\text{m}$. Scale bar, $14 \mu\text{m}$.

electrical conductivity. The thermal conductivity is the sum of the contribution arising from the lattice waves (phonons) K_L and a contribution K_e from the carriers present. All three parameters in the figure of merit vary with the carrier concentration and Z is optimized at around 10^{26}m^{-3} . At this high carrier concentration K_L still accounts for more than 75% of the thermal conductivity².

Interest has grown recently in the use of hot pressing techniques. Apart from facilitating device fabrication, the method of preparation has held out a possibility of reducing the lattice thermal conductivity and hence improving the figure of merit without adversely affecting the other thermoelectric parameters—an effect attributed to phonon scattering at the grain boundaries³.

The role of boundary scattering in limiting the phonon mean free path has long been considered as a low temperature phenomenon, where in the absence of other phonon scattering mechanisms boundary scattering effects are observed as the phonon mean free path approaches the sample dimensions. In solid solutions such as silicon-germanium alloys, however, because a large proportion of the heat is carried by low frequency phonons, boundary scattering effects can manifest themselves in the region of the Debye temperature and above and in principle lead to a reduction in the lattice thermal conductivity.

In thin films of germanium the carrier mean free path is unaffected by boundary scattering down to thicknesses of $0.1 \mu\text{m}$ (ref. 4) and measurements on high density heavily doped hot pressed silicon germanium alloys with grain sized $< 70 \mu\text{m}$ (ref. 5) and $< 10 \mu\text{m}$ (D.M.R. and N. Madini, unpublished data) indicate that the electrical conductivity and Seebeck coefficient differ from single crystal values by less than 5%.

In 1976 a theoretical model was proposed for heavily doped fine grained silicon germanium alloys⁶ but the combination of hot pressing and heavy doping complicates the situation, and in spite of refinements to the model there had been no satisfactory quantitative description of the thermal and electronic behaviour⁷. However, some theoreticians agree that in alloys with a grain size of $< 1 \mu\text{m}$ a reduction in K_L of 20–30% compared with the K_L for single crystal material would be observed.

Table 1 Room temperature properties of investigated Si-Ge alloys

Specimen	K ($\text{W m}^{-1} \text{K}^{-1}$)	S ($\mu\text{V K}^{-1}$)	$\rho = \sigma^{-1}$ ($\times 10^{-5} \Omega\text{m}$)	K_e ($\text{W m}^{-1} \text{K}^{-1}$)	K_L ($\text{W m}^{-1} \text{K}^{-1}$)
Single crystal	5.07	116	0.85	0.76	4.31
$10 < L < 25 \mu\text{m}$	4.93	129	1.05	0.72	4.21
$5 < L < 10 \mu\text{m}$	4.26	133	1.10	0.69	3.57
$L < 5 \mu\text{m}$	3.82	131	1.05	0.72	3.10

* Present address: Division of Chemistry, National Research Council, Ottawa, Canada K1A 0R9.

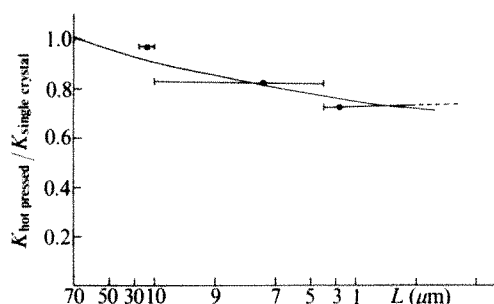


Fig. 2 $K_{\text{hot pressed}}/K_{\text{single crystal}}$ as a function of grain size for heavily-doped silicon-germanium alloys.

We investigated material cut from a single ingot of homogeneous, zone levelled $\text{Si}_{63.5}\text{Ge}_{36.5}$ alloy doped n-type with phosphorus to about $1 \times 10^{26} \text{ m}^{-3}$. A disk shaped sample with suitable dimensions for thermal diffusivity measurement (6 mm diameter $\times$ 1 mm thick) was cut from a slice from the ingot using an ultrasonic cutter. The remainder of the slice was crushed in an agate pestle then ground in an agate vibromill. The fine powder was sieved through a set of British Standard microplate sieves assisted by ultrasonic vibration⁸. Compacts 10 mm in diameter and 3 mm thick were prepared from starting powders possessing a grain size in the range $10 < L < 25 \mu\text{m}$, $5 < L < 10 \mu\text{m}$ and $< 5 \mu\text{m}$ using a vacuum hot pressing technique⁹. The density of the compacts was better than 98% of that of the single crystal starting material. A typical microphotograph of the polished surface of a compact is shown in Fig. 1, and it is apparent that little, if any, grain growth has occurred during compaction. A disk shaped sample was cut from each of the compacts and all four disks annealed at $1,000^\circ\text{C}$ under vacuum for 1 h followed by quenching in cold water to normalize the carrier concentration and hence the value of K_e (ref. 10).

The thermal diffusivity α of the 'single crystal' starting material and of the three compacts was measured using a system based on the laser flash technique¹¹. Measurements on ARMCO iron indicated that the accuracy of the system was better than $\pm 2\%$ at room temperature. Reproducibility of measurements was about the same. Seebeck coefficient data were obtained using a hot probe method (accuracy $\pm 3\%$) and carrier concentration and electrical resistivity values derived from carrier concentration versus Seebeck coefficient¹² and electrical resistivity curves¹³.

The experimental results are displayed in Table 1. Thermal conductivity values were obtained from the measured thermal diffusivity using the relationship $K = C_p \alpha$ where C_p is the specific heat per unit volume at constant pressure and values appropriate to the alloy investigated are taken from the literature¹². The electronic contribution to the thermal conductivity is given by $K_e = \mathcal{L}(k/e)^2 \sigma T$. Where $\mathcal{L}$ depends on the scattering mechanisms of the charge carriers and for the material investigated is ~ 2.9 (ref. 7). The Seebeck coefficient of the three compacts are very similar but all are higher than the 'single crystal' starting material. This increase in Seebeck coefficient and the corresponding increase in electrical resistivity is due to a relatively small loss of dopant during the hot pressing operation. The lattice thermal conductivity decreases with a reduction in grain size and for the compact with a grain size $< 5 \mu\text{m}$ is about 28% less than the 'single crystal' starting material.

Our results are in reasonably good agreement with theory (Fig. 2) and confirm that in fine grained heavily doped silicon germanium alloys phonon scattering at grain boundaries results in a significant reduction in the lattice thermal conductivity. The implication is that thermoelectric generators which employ thermoelements fabricated from fine grained semiconductor alloys should exhibit substantial improvement in performance over those based on 'single crystal' or large grained materials.

We thank the British Gas Corporation for partial support of the project, Professor D. J. Harris for use of departmental facilities, and Dr R. D. Nasby for supplies of alloy.

Received 20 November 1980; accepted 27 February 1981.

1. Bhandari, C. M. & Rowe, D. M. *Contemp. Phys.* **21**, 219-242 (1980).
2. Rowe, D. M. *J. appl. Phys.* **7**, 1843-1846 (1974).
3. Goldsmid, H. J. & Penn, A. W. *Phys. Lett.* **27A**, 523 (1968).
4. Suzuki, K. & Maki, M. *J. appl. Phys. Jap.* **16**, 667 (1977).
5. Nasby, R. D. & Burgess, E. L. *Proc. 7th IECEC*, New York, 130 (1972).
6. Meddins, H. R. & Parrott, J. E. *J. Phys. C: Solid St. Phys.* **9**, 1263 (1976).
7. Bhandari, C. M. & Rowe, D. M. *J. Phys. C: Solid St. Phys.* **11**, 1787-1794 (1978).
8. Mullin, J. W. *Chem. Ind.* 1435-1436 (1971).
9. Bunce, R. W. & Rowe, D. M. *J. Phys. D: Appl. Phys.* **10**, 941-947 (1977).
10. Rowe, D. M. & Savvides, N. *J. Phys. D: Appl. Phys.* **12**, 1613-1619 (1979).
11. Taylor, R. *J. Phys. E: Scient. Instrum.* **13**, 1193-1199 (1980).
12. Dismukes, J. P. et al. *J. appl. Phys.* **35**, 2899-2907 (1964).
13. Rowe, D. M. *J. Phys. D: Appl. Phys.* **8**, 1042-1103 (1975).

Estimate of annual variation of Atlantic Ocean heat transport

Peter J. Lamb

Cooperative Institute for Marine and Atmospheric Studies,
The University of Miami/NOAA, Miami, Florida 33149, USA
and * Climatology Section, Illinois State Water Survey, Champaign,
Illinois 61820, USA

Despite the importance of meridional oceanic heat transport for the understanding of the climate system^{1,2}, most existing hemispheric and ocean basin estimates are necessarily indirect. Annual averages have been obtained for individual basins from sea-air heat exchange results³⁻⁷, and for each hemisphere from measurements of net radiation at the top of the atmosphere and calculations of atmospheric energy transport^{8,9}. The latter approach, when used in conjunction with atmospheric and oceanic energy storage information, has provided monthly mean estimates of the meridional heat transport by the total Northern Hemisphere water body, but not its separate oceans¹⁰. I now present an estimate of the annual variation of meridional heat transport in the Atlantic Ocean, based on large sets of sea-air interaction and subsurface temperature data, which suggests that, with the possible exception of November-December, the heat flux is northward.

Bimonthly averages of the 'vertically and zonally integrated net meridional heat transport' (VZINMHT) within the North and Tropical Atlantic Ocean (70°N - 20°S , including the Gulf of Mexico and Caribbean Sea) were obtained for fixed latitude circle arcs spanning the entire basin (Table 1). The fundamental equation used was that for the heat budget of a local oceanic column,

$$Q_v = Q_{\text{scf}} - Q_t \quad (1)$$

where Q_v is the rate of lateral heat import/export by oceanic motions, Q_{scf} is the rate of net surface heat gain, and Q_t is the column's heat storage change rate. We computed average bimonthly integrals of these terms, with dimensions of Watts, for 10° (extratropics) and 5° (tropics) latitude zones extending across the basin. Those for Q_{scf} were obtained from existing sea-air heat exchange results. The Q_t integrals were estimated using a set of subsurface temperature data acquired and processed for this purpose. These results permitted Q_v to be obtained from equation (1). The zonal Q_v integrals were then converted into VZINMHT estimates (dimensions of Watts) for the zones' bounding latitude circle arcs, by successively integrating southwards from assumed near-zero 70°N VZINMHT boundary conditions computed from ref. 11 (Table 1). The uncertainty in the VZINMHT results given by this procedure naturally increases away from 70°N , and becomes large in the tropics. This is shown in Table 1. Such uncertainty arises from the use of empirical equations to estimate Q_{scf} , and from possible temporal and spatial sampling biases and instrumental inaccuracies in the data sets described below.

* Present address: Climatology Section, Illinois State Water Survey, Urbana, Illinois 61801, USA.

Table 1 Latitudinal variation of bimonthly average vertically and zonally integrated net meridional heat transport (VZINMHT) within the North and Tropical Atlantic Ocean, including the Gulf of Mexico and Caribbean Sea (10^{13} W)

Lat.	January–February		March–April		May–June		July–August		September–October		November–December		Bimonthly uncertainty	Average	Annual uncertainty
	300 m	500 m	300 m	500 m	300 m	500 m	300 m	500 m	300 m	500 m	300 m	500 m			
70° N	+07	+07	+07	+07	+07	+07	+15	+15	+15	+15	+09	+09	5	+10	5
60°	+85	+109	+18	–02	+14	+22	+36	+54	+24	+27	–06	–39	14	+28	8
50°	+126	+171	+57	+32	+29	+40	+59	+78	+50	+73	–15	–86	26	+52	12
40°	+79	+83	+83	+58	+82	+117	+128	+162	+90	+139	–31	–129	42	+73	17
30°	+101	+124	+107	+58	+126	+164	+148	+174	+151	+216	–03	–106	63	+107	24
25°	+101	+136	+126	+81	+113	+147	+169	+185	+157	+230	–05	–117	76	+113	29
20°	+85	+139	+137	+86	+118	+170	+187	+195	+152	+227	–13	–148	90	+114	33
15°	+83	+125	+149	+101	+121	+172	+189	+194	+165	+256	–33	–174	102	+115	37
10°	+90	+128	+197	+157	+75	+118	+180	+191	+162	+246	–45	–181	114	+113	41
5° N	+86	+120	+226	+193	+27	+60	+180	+201	+163	+241	–37	–168	123	+111	44
0°	+63	+79	+196	+164	+18	+61	+180	+203	+211	+295	–77	–209	133	+102	48
5° S	+50	+65	+137	+100	–04	+47	+170	+203	+253	+322	–107	–238	143	+86	51
10°	+59	+77	+138	+108	–13	+35	+156	+194	+257	+319	–121	–256	152	+82	53
15°	+49	+76	+142	+121	–14	+23	+150	+187	+257	+331	–106	–259	161	+83	57
20°	+57	+77	+111	+108	+01	+49	+152	+182	+260	+322	–115	–271	170	+81	60

Northward transports are positive; those for 70° N are boundary conditions computed from Aagaard and Greisman¹¹. Vertical integrations to 300 m and 500 m depths are given separately. The domain treated has the following noncontinental lateral boundaries: Baffin Island east coast, 64° W in Hudson Strait, 66° W in Gulf of St Lawrence, 6° W in Strait of Gibraltar, and 8° E in Skagerrak Strait. The bimonthly uncertainties (10^{13} W) are based on possible errors of 15 W m^{-2} in both Q_{sc} and Q_e (refs 6, 10) and 5×10^{13} W in the 70° N boundary transports. The annual average VZINMHTs listed are computed from annual average Q_{sc} s; they are very similar to the annual means of the bimonthly VZINMHTs cited, due to the annual average Q_e being extremely close to zero. The annual uncertainties (10^{13} W) are based on possible errors of 10 W m^{-2} in the annual average Q_{sc} and 5×10^{13} W in the 70° N boundary transport.

The bimonthly average zonal Q_e integrals were estimated from 233,957 individual temperature soundings, using a 5° square spatial resolution. January–February values were obtained from the difference between the mean heat storage for December–January and February–March, and those for other bimonthly periods were calculated similarly. The temperature soundings used were drawn from the 1967–76 files of a National Oceanographic Data Center (Washington DC) global set which has been fully described by Levitus and Oort¹². Their spatial density is greatest between 30° and 40° N, from which it decreases both northwards and southwards (Table 2). Because of uncertainty about the representativeness of the sub-250 m annual temperature variations¹⁰ (only 54% and 20% of the soundings reach 300 m and 500 m, respectively), Q_e was computed for both the upper 300 m and upper 500 m. Table 1 contains VZINMHT estimates derived from each set of Q_e results.

The bimonthly average zonal Q_{sc} integrals were calculated from the monthly mean sea–air heat exchange results of Bunker (70°–30° N) and Hastenrath and Lamb¹³ (30° N–20° S). Both these studies obtained long-term averages of Q_{sc} from the surface heat budget equation

$$Q_{\text{sc}} = \text{SW}\downarrow - \text{LW}\downarrow - Q_s - Q_e \quad (2)$$

The right-hand terms, which denote the net shortwave ($\text{SW}\downarrow$) and net longwave ($\text{LW}\downarrow$) radiation and the turbulent fluxes of sensible (Q_s) and latent (Q_e) heat, were estimated using conventional empirical equations and several million surface marine meteorological observations obtained from the National Climate Center (Asheville, NC). The data bases, methods, and uncertainties of each investigation have been fully discussed previously^{13–16}. Bunker's results, which are for Marsden (that is 10°) square subareas of varying size¹⁴, were derived from turbulent fluxes computed using the individual observations and variable drag coefficients necessary for the extratropics. Hastenrath and Lamb's results were preferred to Bunker's in the tropics because of the apparent superiority of their $\text{SW}\downarrow$ input¹⁶, and the fact that their 1° square spatial resolution permitted the use of 5° latitude zones in this energy source region for the climate system. Bunker's format required 10° zones.

The most important annual average VZINMHT result in Table 1 is that this flux is directed northwards throughout the study region. It increases from 81×10^{13} W at 20° S to a broad 5°–25° N maximum of 111 – 115×10^{13} W, due largely to substantial enhancement between 5° S and the Equator. There is a pronounced decrease poleward of 30° N, especially between 30° and 40° N. The general pattern of these results is very similar

to that recently obtained by Hastenrath⁶, whereas earlier work found the mean annual VZINMHT to have a northwards maximum near 30° N^{3,5}, weak northwards^{3,4}, or near zero⁵ cross equatorial values, and a southward direction throughout almost all of 0°–20° S (ref. 5). From 40° N southwards, however, Hastenrath's⁶ northwards VZINMHTs are 40–60% larger than the present ones. This difference primarily results from Hastenrath's 40°–50° N Q_{sc} (taken from Budyko¹⁷) being more negative than the Bunker value used here. Budyko¹⁷ failed to detect the warming of the Labrador current between Nova Scotia and Cape Hatteras Identified by Bunker and Worthington¹⁵. The present annual average VZINMHT for 25° N ($+113 \times 10^{13}$ W) is in good agreement with Bryden and Hall's¹⁸ recent direct computation from oceanographic measurements ($+110 \pm 29 \times 10^{13}$ W).

Many of the bimonthly VZINMHT estimates in Table 1 possess considerable uncertainty, which requires that they be treated with some caution. Their most important result is the suggestion that northward transports may prevail throughout almost all of the study region during January–October. The northward flux out of the Northern Tropics seems to be strongest during July–October, when (particularly in September–October) the northwards VZINMHT may also be largest at 20° S and the Equator. This pattern is strongly enhanced for September–October when the storage computations are extended to 500 m. The northwards VZINMHT evidently undergoes only a relatively small reduction in the Northern Tropics during July–October, whereas much more pronounced decreases are indicated for the 40°–50° N (July–August) and 30°–50° N (September–October) zones. May–June, in contrast, seems to be characterized by a northward VZINMHT that is very weak at 20° S and the Equator, increases strongly between 0° and 15° N, approaches that of July–August at 30° N, then decreases uniformly to 50° N.

The most significant March–April result is the pronounced increase in the northward VZINMHT suggested to occur between 5° S and the Equator (Table 1). This converts a moderate northward flux at 20° S into an equatorial one comparable with that of July–October, and seems primarily responsible for the marked 5° S–0° increase in the annual average VZINMHT (Table 1). During this bimonthly period, the 0°–5° S zone experiences both a moderate positive Q_{sc} and pronounced cooling in the upper 300 m (not shown). A progressive decrease of the March–April transport occurs north of 5° N. Table 1 suggests January–February is characterized by a northward VZINMHT that is quite small at 20° S, undergoes a moderate increase to 30° N, and may be stronger than other bimonthly periods at 50° and 60° N. The latter feature, along

Table 2 Number of 1967–76 subsurface temperature soundings available for 10° latitude zones of the Atlantic (including the Gulf of Mexico and Caribbean Sea), and the area of ocean surface (10^{12} m²) in each zone

Latitude zone	Ocean surface area	No. of soundings
60°–70° N	2.94	10,184
50°–60°	4.19	21,913
40°–50°	5.31	45,869
30°–40°	6.95	79,119
20°–30°	8.80	41,233
10°–20°	8.00	21,223
0°–10° N	6.64	6,763
0°–10° S	6.16	4,943
10°–20°	6.12	2,710

The noncontinental lateral boundaries of the domain treated are specified in Table 1.

with the VZINMHT relative minimum at 40° N, arise from the North Atlantic having a maximum negative Q_{stc} between 30° and 40° N at a time when its cooling is strongly concentrated in the 40°–50° N zone (not shown). However, the comparative lack of winter subsurface temperature data for this region (only 44% of summer amount between 40° and 60° N) may render the foregoing results suspect.

The remaining significant feature of Table 1 is that its November–December results are in marked contrast to those for the rest of the year. The VZINMHT is suggested to be directed southwards throughout much or even all of the study region during November–December, and to show a general increase in this direction to a maximum at 10°–20° S. This is particularly indicated by the estimates derived from 500 m Q_i integrals, which have substantial magnitude south of 40° N. Much weaker, and extremely uncertain, southward VZINMHTs were yielded by the 300 m Q_i calculations, especially north of the Equator (Table 1). These November–December results arise from the oceanic cooling, especially when estimated to 500 m, exceeding the negative Q_{stc} north of 10° N (except for 30°–40° N), and from positive Q_{stc} values coinciding with lesser oceanic warming or oceanic cooling south of 10° N (not shown).

Although the November–December results are rather surprising and therefore of potential interest, they should be regarded as highly tentative. In particular, there does not seem to be a fully acceptable rationale for the southward VZINMHT currently available. Furthermore, the higher latitude Q_i and Q_{stc} estimates, on which all VZINMHTs depend, have a poorer data base for the cooler months than the rest of the year. Some relevant statistics were given above. If the November–December VZINMHT is in reality directed northwards rather than southwards, however, the present results probably imply that such northward transport is weaker than in the other bimonthly periods. The November–December results in Table 1 may therefore serve as a challenge to theoreticians and further stimulus for the planned programme to monitor the Atlantic VZINMHT near 25° N (refs 19, 20) directly.

I thank Eric B. Kraus for enabling me to undertake this research, and Stanley A. Changnon Jr for assistance. It was supported by NSF grant ATM 77 28126 to CIMAS, the State of Illinois, and the University of Adelaide Study Leave Programme. The late Andrew Bunker supplied sea–air heat exchange results, and Henry Poor gave computer support. I thank E. Kraus, C. Rooth, J. Merle, S. Hastenrath and W. Wendland for helpful comments.

Received 14 October 1980; accepted 18 February 1981.

- Interdepartmental Committee for Atmospheric Sciences *A United States Climate Program Plan*, 41–44 (Federal Coordinating Council for Science, Engineering, and Technology, Washington DC, 1977).
- US Committee for GARP *Elements of the Research Strategy for the United States Climate Program* 35–39 (National Academy of Sciences, Washington DC, 1978).
- Sverdrup, H. U. *Handbuch der Physik* Vol. 48, 608–670 (Akademie, Berlin, 1957).
- Bryan, K. J. *geophys. Res.* **67**, 3403–3414 (1962).
- Emig, M. J. *geophys. Res.* **72**, 2519–2529 (1967).
- Hastenrath, S. J. *phys. Oceanogr.* **10**, 159–170 (1980).
- Hastenrath, S. & Lamb, P. J. *J. phys. Oceanogr.* **10**, 694–708 (1980).
- Vonder Haar, T. H. & Oort, A. H. *J. phys. Oceanogr.* **3**, 169–172 (1973).
- Trenberth, K. E. *Dynamics Atmos. Oceans* **4**, 57–64 (1979).
- Oort, A. H. & Vonder Haar, T. H. *J. phys. Oceanogr.* **6**, 781–800 (1976).

- Aagaard, K. & Greisman, P. J. *geophys. Res.* **80**, 3821–3827 (1975).
- Levitus, S. & Oort, A. H. *Bull. Am. met. Soc.* **58**, 1270–1284 (1977).
- Hastenrath, S. & Lamb, P. J. *Heat Budget Atlas of the Tropical Atlantic and Eastern Pacific Oceans* (The University of Wisconsin Press, Madison, 1978).
- Bunker, A. F. *Mon. Weath. Rev.* **104**, 1122–1140 (1976).
- Bunker, A. F. & Worthington, L. V. *Bull. Am. met. Soc.* **57**, 670–678 (1976).
- Bunker, A. F. *Bull. Am. met. Soc.* **60**, 131–132 (1979).
- Budyko, M. I. *Atlas of the Heat Balance of the Earth* (Kartfabrika Gosgeoltekhizdata, Leningrad, 1963).
- Bryden, H. L. & Hall, M. M. *Science* **207**, 884–886 (1980).
- National Oceanic and Atmospheric Administration *An Ocean Climate Research Plan*, 25–28 (US Department of Commerce, Washington DC, 1979).
- National Oceanic and Atmospheric Administration *Ocean Heat Transport and Storage*, 5–22 (US Department of Commerce, Washington DC, 1979).

Dissolved state of chromium in seawater

Eiichiro Nakayama, Hiroyuki Tokoro,
Tooru Kuwamoto & Taitiro Fujinaga

Department of Chemistry, Faculty of Science, Kyoto University,
Kyoto 606, Japan

The dissolved state of chromium in seawater has long been studied and in particular which species, trivalent or hexavalent, is predominant. Research results are not, however, consistent, although the method used has usually been the same—using ferric hydroxide as coprecipitation carrier (Table 1^{1–7}). We have studied the coprecipitation behaviour of chromium with ferric hydroxide and other metal hydroxides in the presence of naturally existing inorganic and organic ions to attain an accurate analytical method for determining chromium content, as well as to estimate its behaviour in natural waters^{8,9}. We report here that neither Cr(VI) nor organic species coprecipitate with ferric hydroxide in seawater; Cr(VI) can be quantitatively captured by the coprecipitation with bismuth hydroxide without any pretreatment (such as reduction); and the role of manganese oxide should be considered in addition to the amount of dissolved oxygen to understand the redox system of chromium in natural waters. The inconsistency of the past research may therefore be partly because the presence of organic species was not taken into account and was analysed as either Cr(VI) or Cr(III) or overlooked entirely.

We have developed a method for the fractional quantitative determination of chromium in seawater. A 1.2-l sample of seawater was filtered through a 0.4- μ m nucleopore filter immediately after which the sample was divided into three. To the first and second portions were added respectively ferric hydroxide and bismuth hydroxide precipitates, which had been prepared separately and thoroughly washed with distilled water. In seawater solution with a pH ~8 inorganic Cr(III) only coprecipitates quantitatively with ferric hydroxide, inorganic Cr(III) and Cr(VI) coprecipitate with bismuth hydroxide while the organic species hardly precipitate at all with either hydroxide. The third portion was acidified with hydrochloric acid and heated with ammonium persulphate to decompose any organic matter. The pH of this portion was then adjusted back to 8 with aqueous ammonia and bismuth hydroxide was added. After these portions had been allowed to stand for 12 h, the precipitates were filtered and dissolved in hydrochloric acid. The iron and bismuth was removed by the solvent extraction method using a *p*-xylene solution of 5% tri-isooctylamine, the aqueous phase evaporated to dryness, and the residue dissolved in a small amount of diluted nitric acid. The resulting solution was used for flameless atomic absorption analysis. The total amount of inorganic Cr(III) present was determined from the first portion; the total inorganic Cr(III) and Cr(VI) from the second portion; and the total amount of chromium, including the organic species from the third portion. By performing the experiment with a ⁵¹Cr tracer, it was found that the recovery of chromium was >95% in the coprecipitation procedure and >99% in the other procedures. The procedure was repeated five times with each

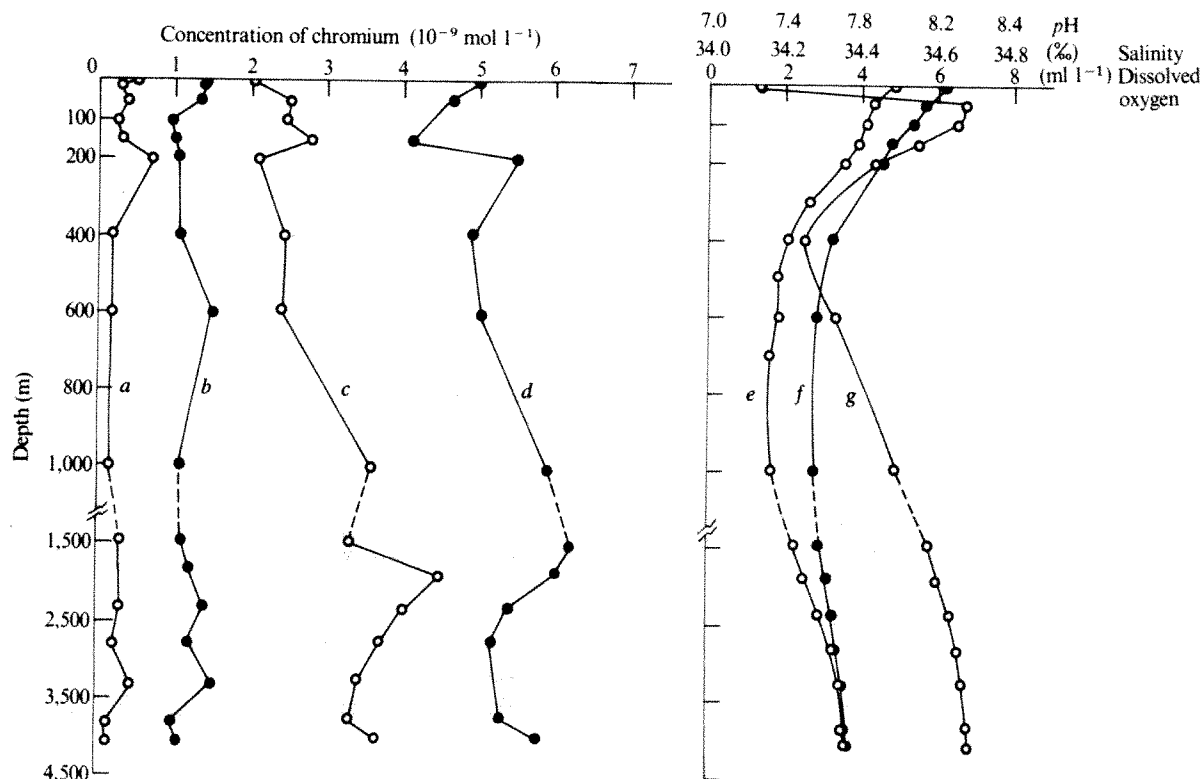


Fig. 1 Vertical distribution of chromium species and some standard components in the Pacific Ocean ($32^{\circ} 19.3' \text{ N}$, $137^{\circ} 33.5' \text{ E}$; depth 4,079 m). a, Particulate chromium; b, dissolved inorganic Cr(III); c, dissolved Cr(VI); d, organic chromium species; e, dissolved oxygen; f, pH; g, salinity.

sample and the standard deviations for the first, second and the third portions were 5.8, 2.4 and 5.1% respectively. The amount of particulate chromium present was determined by analysing the filter paper used for filtrating the sample.

Although there was no particular quality control to minimize chromium contamination, seawater sampling was usually carried out in the direction of the wind and plastic Niskin-type bottles were used to avoid metal contamination in our research vessel.

This method was used to analyse the samples of seawater collected in the adjacent seas of Japan in the summer of 1979. Figures 1 and 2 show typical results from six stations in the Pacific Ocean and four stations in the Japan Sea, respectively. The total average concentration of dissolved chromium is $9.4 \times 10^{-9} \text{ M}$ ($0.48 \mu\text{g l}^{-1}$) in the Pacific Ocean and $8.1 \times 10^{-9} \text{ M}$ ($0.42 \mu\text{g l}^{-1}$) in the Japan Sea: consisting of 10–20% of inorganic Cr(III), 25–40% of Cr(VI) and 45–65% of organic species. The average concentration of particulate chromium is

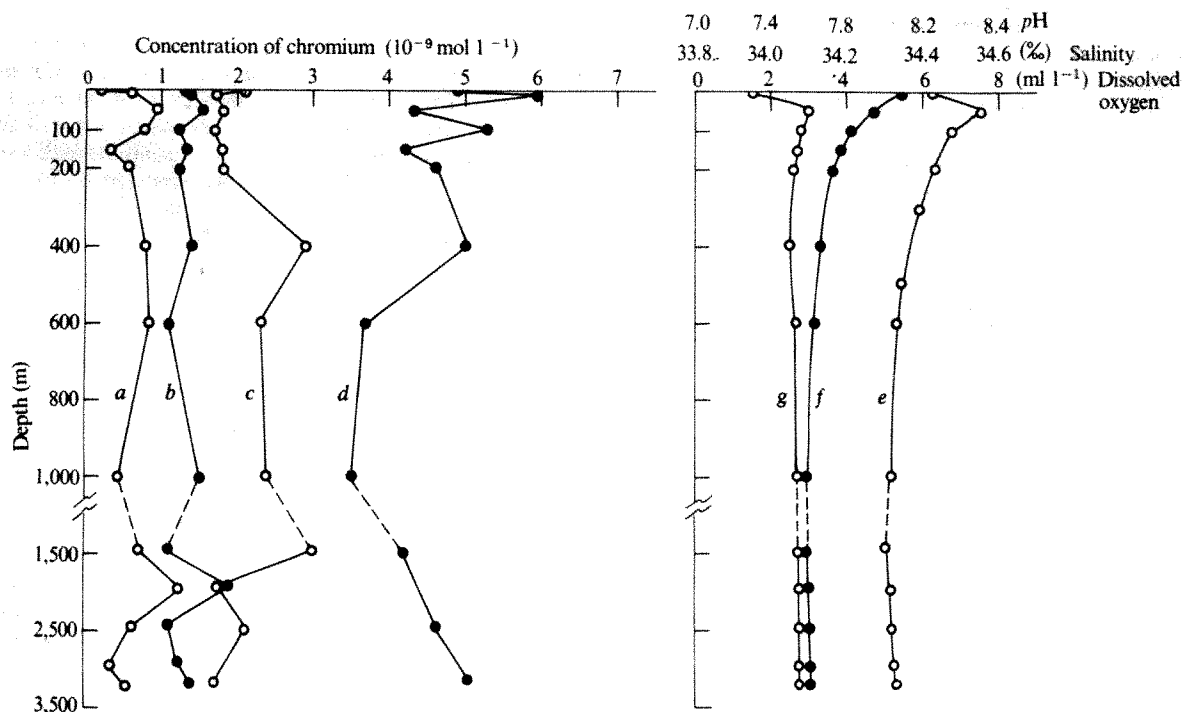


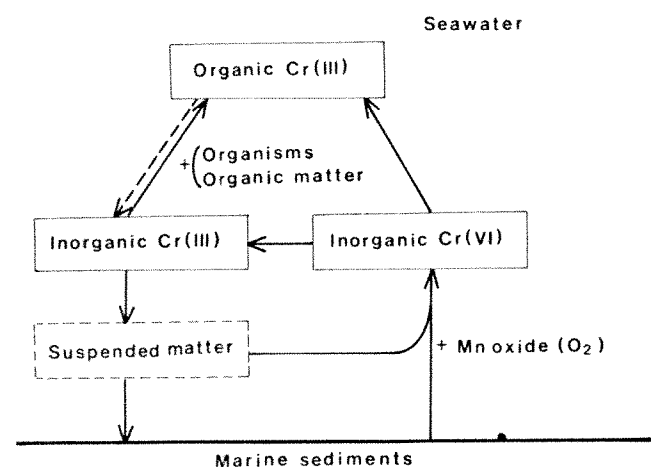
Fig. 2 Vertical distribution of chromium species and some standard components in the Japan Sea ($44^{\circ} 11.9' \text{ N}$, $138^{\circ} 56.4' \text{ E}$; depth 3,447 m). a, Particulate chromium; b, dissolved inorganic Cr(III); c, dissolved Cr(VI); d, organic chromium species; e, dissolved oxygen; f, pH; g, salinity.

Table 1 Literature data on the chromium contents in seawater

Ref.	Locations	Concentration of chromium ($\mu\text{g l}^{-1}$)	Ratio of Cr(VI)/Cr(III)
1	British coastal	Cr(III) 0.46 Cr(VI) 0	Cr(III) only
2, 3	Ligurian Sea	Cr(III) 0.02–0.25 Cr(VI) 0.05–0.36	Cr(VI) > Cr(III)
4	Pacific Ocean	Cr(III) 0.13–0.90 Cr(VI) 0.07–0.27	Cr(VI) < Cr(III)
5	Equatorial region of Pacific Ocean	Cr(III) 0.24–0.52 Cr(VI) 0.03–0.11	Cr(VI) < Cr(III)
6	Northern north Pacific Ocean	Total Cr 0.16–0.95 Cr(VI) 5–70%	Cr(VI) < Cr(III)
7	North-east region of Pacific Ocean	Cr(III) 0.005 Cr(VI) 0.15	Cr(VI) > Cr(III)

$3.6 \times 10^{-10} \text{ M}$ ($0.07 \mu\text{g l}^{-1}$) in the Pacific Ocean and $4.2 \times 10^{-10} \text{ M}$ ($0.08 \mu\text{g l}^{-1}$) in the Japan Sea. Figures 1 and 2 show that the vertical distribution of Cr(VI) differs between the two Seas. In the Pacific Ocean there is a remarkable increase at a depth of 1,000 m and below while in the Japan Sea there is a decrease. The ratio of Cr(VI) to inorganic Cr(III) averages 2.7 in the Pacific Ocean while in the Japan Sea it averages 1.8. This difference cannot be explained by thermodynamical equilibrium and application of the oceanographical data, such as pH and dissolved oxygen¹⁰, because the two seas are completely oxygenated. Hence we propose a model for the circulation of chromium in seawater as shown in Fig. 3. This model was supported by an experiment based on the geochemical fact that in manganese nodules the chromium content is extremely low compared with that in the marine sediments composed of clay minerals¹¹. Thus, when the seawater solution containing 10^{-5} M of Cr(III) was bubbled with air for >300 h no hexavalent chromium was detected. However, when about 50 mg l^{-1} of powdered manganese oxide (such as synthesized manganite and manganese nodules) were added, 10% of Cr(III) was oxidized within 100 h. Organic complexes of Cr(III) (such as chromic citrate) were not oxidized within 100 h even if excess manganese oxide was present. From this model, the behaviour of chromium in seawater can be explained as follows. When inorganic Cr(III) is adsorbed by suspended particles it is oxidized by the catalytic action of manganese oxide during sedimentation or in the marine sediments. Direct oxidation with dissolved oxygen only is not appreciable. Cr(VI) is reduced by being taken into organisms or by reductive organic matter. When organic species form, oxidation ceases.

Thus we interpret the difference of distribution of Cr(VI) between the two Seas as being due to the fact that the marine sediments contain abundant manganese oxide in the Pacific Ocean because they are oxygenated down to a considerable depth, whereas the sediments in the Japan Sea, except for a thin oxygenated layer at the surface reduce manganese oxide to Mn(II) (ref. 12). Accordingly, Cr(VI) is supplied to the water

**Fig. 3** A model for the circulation of chromium in seawater.

column in the Pacific Ocean, but not to the Japan Sea. In addition, the Cr(VI) supplied from suspended particles does not seriously effect its vertical distribution because their chromium content is very low.

We thank members of the staff of the Ocean Research Institute, University of Tokyo for their cooperation in the collection of samples.

Received 25 September 1980; accepted 24 February 1981.

1. Chuecas, L. & Riley, J. P. *Analyt. chim. Acta* **35**, 240–246 (1966).
2. Fukai, R. *Nature* **213**, 901 (1967).
3. Fukai, R. & Vas, D. J. *oceanogr. Soc. Jap.* **23**, 298–302 (1967).
4. Kuwamoto, T. & Murai, S. *Preliminary Rep. Hakuho-maru Cruise KH-68-4*, 72 (Ocean Research Institute, University of Tokyo, 1970).
5. Grimaud, D. & Michard, G. *Mar. Chem.* **2**, 229–237 (1974).
6. Yamamoto, T., Kadowaki, S. & Carpenter, J. H. *Geochem. J.* **8**, 123–133 (1974).
7. Cranston, R. E. & Murray, J. M. *Analyt. Chim. Acta* **99**, 275–282 (1978).
8. Fujinaga, T., Kuwamoto, T., Murai, S., Kihara, S. & Nakayama, E. *Nippon Kagaku Zasshi* **92**, 339–344 (1971).
9. Fujinaga, T., Kuwamoto, T., Nakayama, E. & Tsurubo, S. *Abstr. Papers ACS/CSJ Chem. Congr. Envir. Chem.* No. 100 (Honolulu, Hawaii, 1979).
10. Elderfield, H. *Earth planet. Sci. Lett.* **9**, 10–16 (1970).
11. Turekian, K. K. *Chemical Oceanography* Vol. 2, 81–125 (Academic, New York, 1955).
12. Masuzawa, T. *Abstr. oceanogr. Soc. Jap.*, 148–149 (1978).

Effect of irregular fluctuations in Antarctic precipitation on global sea level

J. Oerlemans

Institute of Meteorology and Oceanography, University of Utrecht, Postbus 80.005, 3508 TA Utrecht, The Netherlands

One of the reasons for the continuing interest in the global sea level is that secular variations may be caused by climatic changes. Such a change could, for example, be an atmospheric warming due to CO_2 accumulation. Changes in the amount of ice in the major ice sheets will be reflected in secular variations of sea level; it has, for example, been suggested that ice-shelf thinning may change the drainage of parts of the Antarctic Ice Sheet¹. Attempts to monitor climatic change by measuring global sea level will, however, be complicated by random fluctuations of the ice volume in the major ice sheets, themselves the consequence of random variations in the ice accumulation rate. Precipitation rates are highly variable, and this also applies to Antarctica², which stores most of the continental ice mass. By means of a simple model for ice flow in the Antarctic, together with proxy data on precipitation variability derived from ice cores, I show that long-term sea-level variations with a standard deviation of roughly 5 cm are to be expected on this account. This 'climatic noise' is comparable in magnitude with many of the secular effects now being sought.

Any changes in the volume of the Antarctic Ice sheet are damped by a mechanism involving several drainage systems³. In these systems ice is discharged to the ocean by horizontal pressure gradients resulting from differences in elevation of the ice surface. If the ice thickness of a drainage system is increased (without changing the mass balance—the yearly gain of ice mass at the surface), the mean pressure gradient from the interior of the drainage system to the ocean will become larger. The subsequent increase in ice-mass discharge will try to restore the original ice thickness. In the following, we will measure the strength of this damping by the relaxation time, t_R , which indicates how fast the system can return to equilibrium.

Once this time scale is known, the effect of random fluctuations in the mass balance on the volume of the Antarctic Ice Sheet can be estimated with a simple stochastic model. A framework for the application of stochastic models to climate problems has been presented by Hasselmann⁴. Here I will use a simple model, consisting of a rate equation for ice thickness with a linear damping term and an additive stochastic term to represent the fluctuations in the ice accumulation rate.

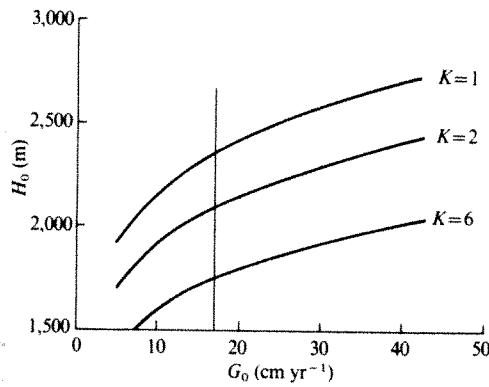


Fig. 1 Steady-state ice thickness H_0 , as a function of the mass balance G_0 , for three values of the flow parameter K ($\text{m}^{-3/2} \text{yr}^{-1}$). The vertical line indicates the present mass balance over Antarctica.

To obtain a value for t_R , we can turn to a schematic physical description of the dynamics of large ice streams. As proposed by Nye⁵, we use a flow law for the vertically integrated horizontal ice velocity (denoted by u), which should account for both internal deformation and (moderate) basal sliding. It reads

$$u = k\tau_b^m \quad (1)$$

where τ_b is the basal shear stress and k and m are constants. The basal shear stress is proportional to the ice thickness, H , times the slope of the ice surface. For simplicity, we assume that the bedrock is flat. After some manipulations, a one-dimensional ice-sheet model based on the conservation of ice mass, can be formulated as follows

$$\frac{\partial H}{\partial t} = \frac{\partial}{\partial x} \left[D \frac{\partial H}{\partial x} \right] + G(x) \quad (2)$$

where

$$D = kH^{m+1} \left| \frac{\partial H}{\partial x} \right|^{m-1}.$$

The x axis is oriented in the direction of the mean ice velocity in the drainage system, that is, from the interior of the Antarctic continent to the ocean. The mass balance is denoted by $G(x)$. I have used the model formulated by equation (2) in some other studies. (See ref. 6 for details.) Current values for m are within the 2–3.5 range. Here we adopt a value of 2.5. The flow parameter, k , will be chosen such that the present model produces the correct order of magnitude of ice thickness in a typical Antarctic drainage system.

To make the model more tractable, we will now make it zero-dimensional by inserting in equation (2) typical values of the quantities involved. Let the typical ice thickness be H_0 and the typical horizontal scale for ice thickness variations L_0 . To a first approximation, the dynamics of the drainage system can be described by

$$\frac{dH_0}{dt} = -KH_0^{2m+1}L_0^{-(m+1)} + G_0 \quad (3)$$

The constant, K , is not necessarily the same as k . The steady-state ice thickness can easily be found by setting the dH_0/dt to zero. The resulting relation between H_0 and G_0 is displayed in Fig. 1. L_0 has been set to 1,000 km. Curves are shown for $K = 1, 2$ and $6 \text{ m}^{-3/2} \text{yr}^{-1}$, corresponding to $H_0 = 2,311, 2,055$ and $1,712 \text{ m}$ respectively, for $G_0 = 0.17 \text{ m}$ of ice per yr (which is about the average mass balance of the Antarctic Ice Sheet).

To fit the description of ice-flow dynamics given above into a simple linear stochastic model, equation (3) has to be linearized around some reference state. For this purpose, it is natural to use the present state (although we cannot be sure that at present the Antarctic Ice Sheet is close to equilibrium), which will be

denoted by asterisks. We find

$$\frac{dH'_0}{dt} = -K(2m+1)H_0^{*2m}L_0^{-(m+1)}H'_0 + G'_0 \quad (4)$$

A prime denotes deviation from the reference state. According to equation (4), the relaxation time t_R equals $[K(2m+1)H_0^{*2m}L_0^{-(m+1)}]^{-1}$. Values of t_R corresponding to $H_0^* = 2,353, 2,097$ and $1,746 \text{ m}$ are 2,311, 2,055 and 1,712 yr, respectively. Apparently, t_R is not very sensitive to the particular reference state we choose.

At this point, we interpret $G'_0(t)$ as a stochastic process, representing random fluctuations in the mass balance. The quantity H'_0 is assumed to represent deviations of the mean thickness of the Antarctic Ice Sheet (note that this implies that the various drainage systems have the same dynamical characteristics, which seems a reasonable approximation for this analysis). To keep things as simple as possible, we assume that $G'_0(t)$ is a white-noise process, which implies that the process $H'_0(t)$ has a spectrum proportional to $1/(1+\omega^2 t_R^2)$ (red noise). The total variance of $H'_0(t)$ equals $t_R S_G$, where S_G is the (constant) spectral density of the white noise⁴. In practice, a reliable estimate of S_G is best obtained from mass balance measurements that are averaged over some time. If the averaging time is t , the following relation results:

$$\sigma_{H_0} = (t_R)^{1/2} \sigma_G \quad (5)$$

Here, σ denotes standard deviation and the tilde indicates that the mass balance is averaged over t yr. Choosing $t_R = 2,100 \text{ yr}$ (intermediate curve in Fig. 1) and $t_R = 30 \text{ yr}$, we have $\sigma_{H_0} = 250\sigma_G$.

At this point an estimate of σ_G is needed. Figure 2 shows some proxy data on mass balance variations that were readily available and permit an order-of-magnitude estimate of σ_G . The records are inferred from stratigraphical studies at three locations on the Greenland Ice Sheet⁷ and at the South Pole⁸. The data shown suggest a relative standard deviation for 30-yr mean values of roughly 7%. For the Antarctic Ice Sheet with a mean mass balance of 0.17 m ice per yr, this gives $\sigma_G = 0.012 \text{ m}$ ice yr^{-1} . It cannot be expected that precipitation anomalies are uniform over Antarctica. To correct for this we need to know into how many 'independent' regions Antarctica can be divided. At present, no data are available on the typical horizontal scale of accumulation anomalies over the Antarctic Ice Sheet. We therefore arbitrarily assume that five regions exist where

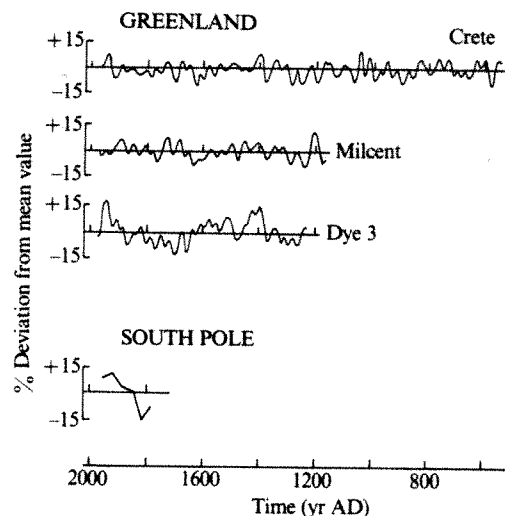


Fig. 2 Time series showing variations in ice accumulation rates, as inferred from stratigraphical studies on the ice of Greenland and Antarctica. The three upper curves are from ref. 7. Variability on time scales smaller than 30 yr has been removed from these curves. The lower curve shows 30-yr mean values of accumulation at the South Pole, computed from data given in ref. 8. Variability is expressed as relative deviation from the mean value.

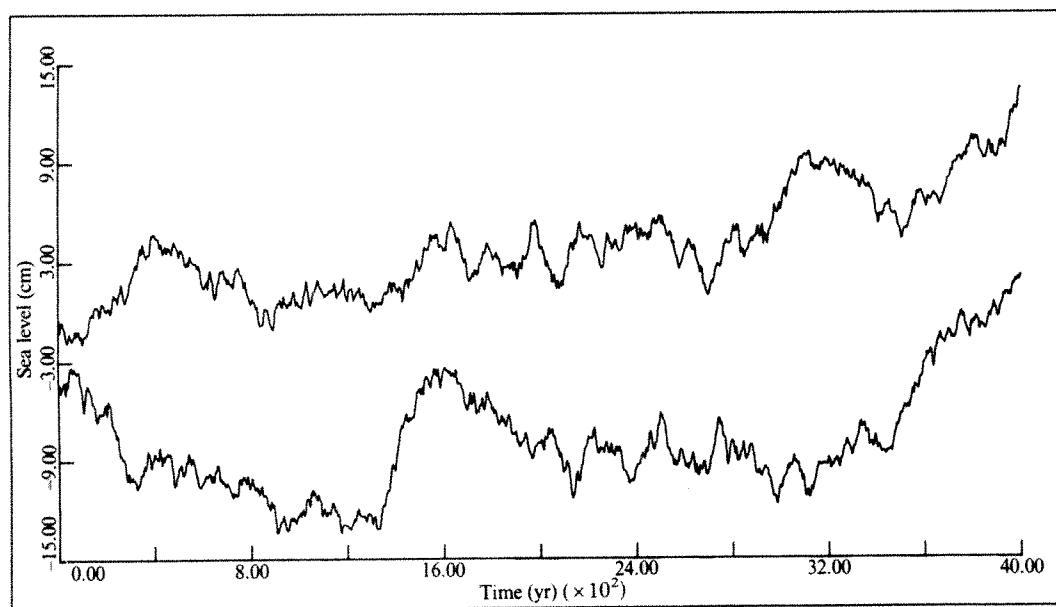


Fig. 3 Two sea-level records generated by the stochastic model, with a length of 4,000 yr. Values of the constants used: $t_R = 2,100$ yr, $\sigma_G = 0.85$ cm ice per yr.

snowfall fluctuations are uncorrelated. The value of σ_G should then be divided by $\sqrt{5}$, which gives $\sigma_G = 0.0054$ m ice yr^{-1} . The resulting value of σ_{H_0} is about 1.35 m. As the area of the Antarctic Ice Sheet is $\sim 4\%$ of that of the world ocean, the corresponding standard deviation of the sea level amounts to ~ 5 cm. If we would choose the spatial scale of precipitation much smaller, the value of σ_{H_0} would also be smaller. For example, if nine independent regions were present, σ_{H_0} would be about 1 m.

To show more clearly what happens, Fig. 3 provides two examples of sea-level records generated by equation (4). The process $G_0(t)$ was simulated by a random number generator, such that $\sigma_G = 0.008$ m ice yr^{-1} . The records, both spanning, 4,000 yr, show that the major part of the variability is contained in the longer time scales. This is to be expected, because the relaxation time of the ice sheet is of the order of a few thousand years.

The foregoing results clearly demonstrate that stochastic forcing of the Antarctic Ice Sheet by fluctuations in snow accumulation may contribute significantly to variations in global sea level. As revealed in Fig. 3, a sea-level trend of 4 cm in 100 yr can easily be due to the process analysed in this study. Even if local sea-level changes are considered, this is by no means negligible. Other mechanisms creating long-term sea-level variability (like changes in atmospheric wind and pressure patterns, shifts in the oceanic circulation regimes, eustatic effects) lead to changes with the same order of magnitude⁹.

It should be stressed that I have presented here an order-of-magnitude estimate. In the foregoing analysis several uncertainties exist. First, the estimate of σ_G is very crude. A critical survey of all data on accumulation over the Antarctic continent is needed to derive a typical horizontal scale for accumulation anomalies. Another point concerns the very schematic treatment of the dynamics of the Antarctic Ice Sheet. The geometry of Antarctica is rather complicated (large variations in bedrock elevation), so modelling of the ice flow by a zero-dimensional model can only give a crude approximation to reality. The assumption of constant L_0 is a poor approximation for regions where the Antarctic continent is not bounded by deep ocean but by a Antarctic shelf, which makes it possible for the ice sheet to expand horizontally.

In view of these points, perhaps the vertical axis in Fig. 3 should be doubled, or halved. In spite of this uncertainty, however, this study reveals that a substantial part of global sea-level variability may be caused by random forcing of the Antarctic Ice Sheet.

Received 3 December 1980; accepted 18 February 1981.

1. Thomas, R. H., Sanderson, T. J. O. & Rose, K. E. *Nature* **277**, 355–358 (1979).
2. Schwerdtfeger, W. in *World Survey of Climatology* Vol. 14 (ed. Landsberg, H. E.) 253–355 (1970).
3. Giovinetto, M. B. in *Antarctic Snow and Ice Studies*, Antarctic Res. Ser. 2 (ed. Mellor, M.) 127–155 (American Geophysical Union, 1964).
4. Hasselmann, K. *Tellus* **28**, 473–485 (1976).
5. Nye, J. F. *J. Glaciol.* **3**, 493–507 (1959).
6. Oerlemans, J. *Tellus* **33**, 1–12 (1981).
7. Reeh, N. et al. *J. Glaciol.* **20**, 27–30 (1978).
8. Giovinetto, M. B. & Schwerdtfeger, W. *Arch. Meteorol. Geophys. Bioklimatol.*, Ser. A **15**, 228–250 (1966).
9. Lisitzin, E. *Sea-level Changes* (Elsevier, Amsterdam, 1974).

The Caledonian nappes of eastern North Greenland

John Malcolm Hurst* & William Stuart McKerrow†

* Grønlands Geologiske Undersøgelse, Øster Voldgade 10, 1350 Copenhagen K, Denmark

† Department of Geology and Mineralogy, Parks Road, Oxford OX1 3PR, UK

Reconnaissance work in Kronprins Christian Land^{1–7} indicated the presence of one main nappe of late Proterozoic sediments^{2,3} thrust over Lower Palaeozoic platform carbonates similar to Peary Land^{7,8}. Our 2-month helicopter-supported work in 1980 now indicates the presence of three thrust-bounded nappes. The Finderup Land Nappes, which occur in northern Kronprins Christian Land, contain a late Precambrian early Cambrian sequence in thrust contact above Silurian carbonates. The younger Vandredalen Nappe in thrust contact above the Finderup Land Nappes and Silurian carbonates occurs throughout Kronprins Christian Land and contains a thick sequence of late Proterozoic clastics and carbonates. In southern Kronprins Christian Land the Sæfaxi Elv Nappe only contains Ordovician–Silurian carbonates and is in thrust contact below Vandredalen Nappe. The nappes were involved in westward-directed thrusting (>150 km) and are probably a result of continental collision when the northern part of Iapetus Ocean closed in the late Silurian.

On the east coast of Kronprins Christian Land (Fig. 1), a narrow belt of gneisses crops out to the east of Proterozoic sandstones which have been referred previously to the Thule Group^{5,6,9}. Both are probably equivalent to the Proterozoic Independence Fjord Group west of Danmark Fjord¹⁰ (Fig. 1).

The Proterozoic sandstones are intruded by numerous dolerite dykes and sills. West of Danmark Fjord (Fig. 1) an intrusive granophyre in the Independence Fjord Group has yielded a whole-rock Rb-Sr isochron age of $1,230 \pm 25$ Myr (ref. 11) and radiometric dates (minimum age) of interbedded clays give a 1,380 Myr age¹². The Proterozoic sandstones of Kronprins Christian Land are probably of a similar age ($>1,000$ Myr).

The Proterozoic sandstones in Kronprins Christian Land are locally overlain by a conglomerate³ consisting of large (10–100 cm) angular to rounded boulders of Proterozoic sandstone and dolerite. This conglomerate is followed at Hjørnegletscher (E, Fig. 1) by a thin sequence of shallow marine clastics and a stromatolitic dolomite identical to the late Precambrian Fyns Sø Formation of the area west of Danmark Fjord; but at Marmorvigen (C, Fig. 1) the Fyns Sø Formation rests directly on the conglomerate.

The Proterozoic sandstones in Kronprins Christian Land are deformed by strong Caledonian movements (H. F. Jepsen and F. Kalsbeek, personal communication). In Ingolf Fjord the Proterozoic sandstones and the overlying late Precambrian shallow marine clastic sediments are involved in a series of broad anticlines and narrow inclined synclines, showing overriding movement to the west. South of Ingolf Fjord (Fig. 1) these folds are replaced by westward-directed thrusts of Proterozoic sandstone over Fyns Sø Formation. Although the Proterozoic sandstones are in part allochthonous in southern Kronprins Christian Land, they have not moved far, as the thrusts in the south pass laterally into folds in the north.

To the west of the Proterozoic sandstones three large thrust-bounded nappes (Fig. 1) (or groups of nappes) with distinctive

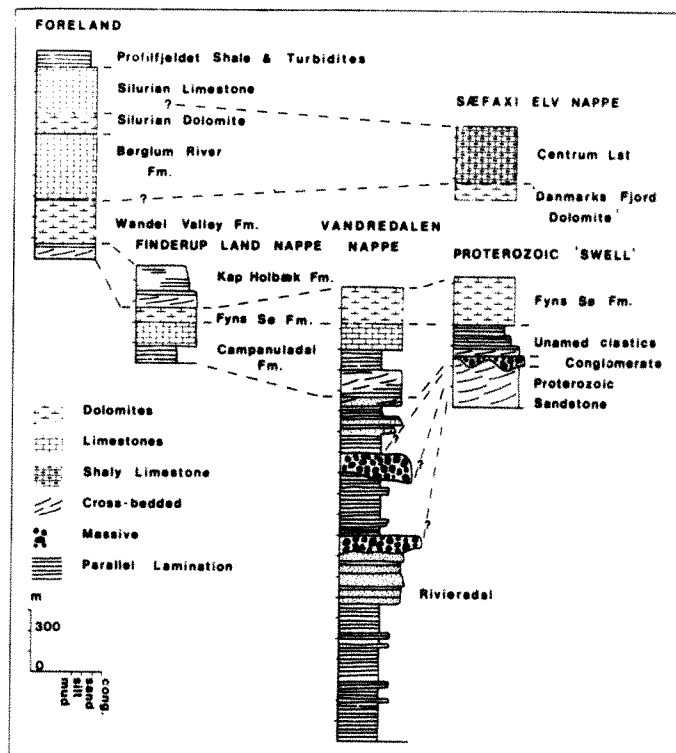


Fig. 2 Stratigraphy and facies distribution of the autochthonous (Foreland⁷ and Proterozoic 'Swell') and allochthonous (Finderup Land Nappes, Vandredalen Nappe and Sæfaxi Elv Nappe) sediments.

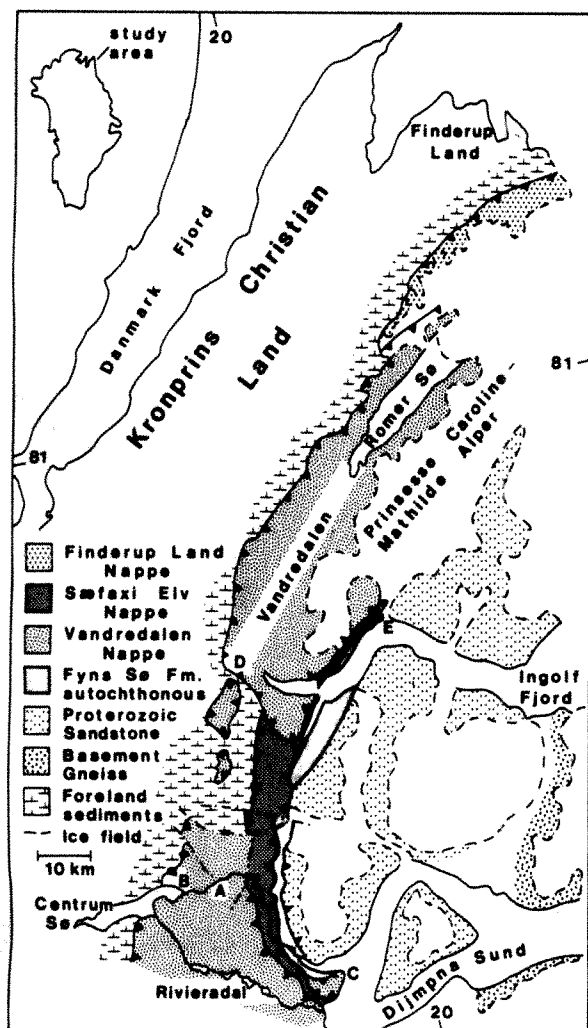


Fig. 1 Simplified geological map of Kronprins Christian Land, North Greenland.

stratigraphic sequences are present in Kronprins Christian Land: (1) The Finderup Land Nappes in the north (Fig. 1). These contain sequences of stromatolitic limestone and clastics of the Campanuladal Formation and deformed stromatolitic dolomite of the Fyns Sø Formation overlain by sandstones with *Skolithos* and bioturbated siltstones of the Cambrian Kap Holbæk Formation. (2) The Vandredalen Nappe of central and southern Kronprins Christian Land (Fig. 1). This consists of a thick sequence of turbidites (Rivieradal sandstones^{2,3}) overlain by sands, silts, variegated shales, silts and red limestones (in part Campanuladal Formation) and stromatolitic dolomites (Fyns Sø Formation). This stratigraphic sequence is well exposed on Drømmebjerg (B, Fig. 1), on Nøglefjeldet (D, Fig. 1) and other hills west of Vandredalen. Basal turbidites of the Rivieradal unit also occur on a higher nappe at Marmorvigen (C, Fig. 1) and similar silty turbidites may be on separate nappes in northern Kronprins Christian Land. (3) The Sæfaxi Elv Nappe in south-east Kronprins Christian Land (Fig. 1). This contains a sequence of dolomites with sandstone fissures which have been referred to the Danmarks Fjord dolomite^{2,3}, overlain by Ordovician to ?Silurian black limestones and limy shales of the Centrum Limestone unit.

In addition to these large thrust-bounded nappes, several complex imbricate zones are exposed along their eastern and western margins. To the south-east, near Marmorvigen and along the south-west side of the Sæfaxi Elv (A, Fig. 1), the Vandredalen Nappe rests on the Sæfaxi Elv Nappe and both are thrust over a sequence of Fyns Sø Formation and the conglomerate above the Proterozoic sandstones and dolerites.

The Finderup Land Nappes occur to the north of a large ice cap (Fig. 1) so that its relationship to the Vandredalen Nappe is not clearly exposed. However, the late Precambrian and Cambrian Fyns Sø Formation and Kap Holbæk Formation of the Finderup Land Nappes appear to underlie the Vandredalen Nappe, and this is borne out by the fact that the beds of this nappe are directly in thrust contact with the folded Silurian platform carbonates in Finderup Land.

To the west, the nappes are in thrust contact with the folded sequence of Ordovician and Silurian platform carbonates and

Silurian trough clastics⁷. In the vicinity of the thrusts these Palaeozoic rocks are locally overturned towards the west. If the strike of the fold axes is a direct response to the direction of thrusting, the movement of the Vandredalen Nappe was between 260° and 280°, while in the north the Funderup Land Nappes seem to have travelled towards 310°. To the west and north-west of the nappes, the Palaeozoic beds show asymmetrical folds with steep or overturned north-west or west facing limbs to the anticlines. This disturbed zone stretches for up to 40 km west of the present exposures of the nappes⁷ and, if it is due to deformation by the overriding nappes, it is possibly an indication that the nappes may have formerly extended to Danmark Fjord. Although some steeply inclined reversed faults occur in this region of disturbed Palaeozoic strata, none of the beds has been transported far; they are essentially autochthonous⁷. To the west of Danmark Fjord the same Palaeozoic platform strata are flat-lying and rest undisturbed on earlier Precambrian sediments.

The stratigraphic sequences (Fig. 2) of the large nappes and of the Proterozoic sandstones to the east can only be correlated by: (1) The presence of the stromatolitic dolomites of the Fyns Sø Formation which are known to be late Precambrian on the platform to the west. (2) The presence of *Skolithos* and bioturbated silts which suggests that the uppermost beds of the Funderup Land Nappes are Lower Palaeozoic, and perhaps equivalent to the Lower Cambrian Kap Holbæk Formation of Danmark Fjord. (3) The presence of Ordovician³ and ?Silurian fossils in the Centrum Limestone of the Sæfaxi Elv Nappe. (4) The similarities between the basal lenticular conglomerates resting on the Proterozoic sandstones and those within the Rivieradal unit.

Although none of these individually provides definite age correlations we believe that together they indicate the correlations suggested in Fig. 2.

The Funderup Land Nappes both contain sequences of laminated and cross-bedded sandstones of the Kap Holbæk Formation (with *Skolithos*) above the stromatolitic dolomites of the Fyns Sø Formation. The derivation of these nappes from the south-east suggests that early sliding occurred from an area of uplift to the east of southern Kronprins Christian Land. The folding associated with the northwesterly emplacement of the Funderup Land Nappes has resulted in an extensive area (including both allochthonous and autochthonous rocks) of northern Kronprins Christian Land having strikes at variance with the rest of East Greenland and Kronprins Christian Land in particular.

We agree with Fränkl^{2,3} that the other large nappes are derived from areas east of Kronprins Christian Land. The earlier Sæfaxi Elv Nappe mainly consists of a deep water facies of the Ordovician to ?Silurian Centrum Limestone, containing about 400 m of thin-bedded black limestones and shales. The Sæfaxi Elv Nappe which underlies the Vandredalen Nappe, consists of entirely younger strata. Both nappes have apparently been derived from a source area east of the present day Proterozoic sandstone outcrop (Fig. 1), and both have deeper water facies than equivalent beds on the platform to the west.

The oldest beds on the Vandredalen Nappe are here all assigned to the Rivieradal unit of Fränkl^{2,3}. He named several phyllite, greywacke and sandstone units and considered several nappes were present, but we believe that all the late Precambrian beds below the Campanuladal Formation on this nappe (Fig. 2) are part of a single turbidite and mud sequence on a single nappe which contains several areas (up to ~20 km²) of low-angle isoclinal folds. Because of its tectonic deformation it is not possible to estimate accurate thicknesses, but the lower part of the sequence comprises at least 1,000 m of mudstones (now slates and phyllites) and thin-bedded fine sand turbidites (some rich in calcareous detritus); and the upper part (about 1,500 m) includes thin- to thick-bedded, medium to coarse sand turbidites in addition to several mudstone units and two conglomerate horizons with rounded Proterozoic sandstone and dolerite clasts. It was one of these mass flow resedimented

conglomerates that Fränkl^{2,3} wrongly interpreted as a tillite. These turbidite sequences are followed by several hundred metres of shelf clastics, ~600 m of quartz-rich sandstones, multicoloured shales and red limestones of the Campanuladal Formation and finally by about 200 m of stromatolitic dolomites of the Fyns Sø Formation (Fig. 2).

The presence of Fyns Sø Formation stratigraphically above the Proterozoic sandstones in eastern Kronprins Christian Land (Fig. 2) implies that the Vandredalen Nappe, which also includes the Fyns Sø Formation, must have been derived from east of the Proterozoic sandstone outcrops in the Prinsesse Caroline-Mathilde Alper (Fig. 1). Fränkl^{2,3} first suggested that the sediments contained on this nappe came from the east. It seems probable that the allochthonous thick sedimentary sequences were deposited east of a swell formed by the Proterozoic sandstones which were covered only by thin late Precambrian strata below the Fyns Sø Formation. The conglomerate on the Proterozoic sandstone swell indicates local uplift of the Proterozoic sandstones and this swell may have been the source of the similar rounded clasts in the resedimented conglomerates in the upper parts of the Rivieradal unit.

If, as seems probable, the Funderup Land, Vandredalen and Sæfaxi Elv Nappes are all derived from the margins of the Iapetus Ocean¹³, they would probably have originated east of the gneiss outcrops in eastern Kronprins Christian Land. A southeasterly source for the Funderup Land Nappes would imply a total transport distance of over 150 km, whereas an easterly source for the Vandredalen and Sæfaxi Elv Nappes suggests transport of over 100 km.

The youngest age for the Silurian Profilfjeldet Shale and Turbidites, which were deformed by the overriding nappes, gives a maximum age limit for the Caledonian Orogeny in East Greenland. A precise minimum age of the thrusting in Kronprins Christian Land cannot be established as no late Silurian or Devonian rocks are known. The Profilfjeldet Shale and Turbidites in Kronprins Christian Land may extend into late Wenlock time; the same turbidites in Peary Land extend into the middle Wenlock¹⁴⁻¹⁶. The Profilfjeldet Shale and Turbidites of Kronprins Christian Land and the Peary Land turbidites are mainly derived from the east. In both regions the turbidites succeed a prolonged episode of Ordovician and Lower Silurian platform carbonates¹⁶.

The Caledonian Orogeny in Norway occurs at the same time as in eastern North Greenland. Wenlock rocks are involved in eastward-directed thrusting near Trondheim¹⁷ and Pridoli (Downtonian) conglomerates near Bergen are post-orogenic¹⁸. It seems likely that the Caledonian Orogeny in both Greenland and Norway was due to continental collision in the late Wenlock and Ludlow, closing these northern parts of the Iapetus Ocean. It is also possible that just before collision some local areas of uplift provided a source area for the extensive Silurian turbidites of North Greenland¹⁶ in the late Llandovery and early Wenlock.

We thank N. Henriksen, H. F. Jepsen, F. Kalsbeek and J. S. Peel for critically reading the manuscript. This letter is published with the permission of the director of the Geological Survey of Greenland.

Received 18 December 1980; accepted 18 February 1981.

- Nielsen, E. *Meddr Grønland* **126**, 1-35 (1941).
- Fränkl, E. *Meddr Grønland* **116**, 1-85 (1954).
- Fränkl, E. *Meddr Grønland* **103**, 1-35 (1955).
- Adams, P. J. & Cowie, J. W. *Meddr Grønland* **111**, 1-24 (1953).
- Haller, J. *Meddr Grønland* **171**, 1-286 (1950).
- Haller, J. *Geology of the East Greenland Caledonides* (Interscience, London, 1971).
- Peel, J. S. *Rapp. Grøn. geol. Unders.* **99**, 61-72 (1980).
- Christie, R. L. & Peel, J. S. *Rapp. Grøn. geol. Unders.* **82**, 1-48 (1977).
- Troelsen, J. C. *Meddr Grønland* **149**, 1-29 (1949).
- Collinson, J. D. *Rapp. Grøn. geol. Unders.* **99**, 7-23 (1980).
- Jepsen, H. F. & Kalsbeek, F. *Rapp. Grøn. geol. Unders.* **88**, 11-14 (1979).
- Larsen, O. & Graff-Pedersen, P. *Rapp. Grøn. geol. Unders.* **99**, 111-118 (1980).
- Harland, W. B. & Gayer, R. A. *Geol. Mag.* **109**, 289-314 (1972).
- Hurst, J. M. & Surlyk, F. *Rapp. Grøn. geol. Unders.* **99**, 73-78 (1980).
- Lane, P. D. *Palaeontology* **15**, 336-364 (1972).
- Surlyk, F., Hurst, J. M. & Bjerreskov, M. *Nature* **286**, 800-803 (1980).
- Gee, D. *Tectonophysics* **47**, 393-419 (1978).
- Nicholson, R. in Harris, A. L., Holland, C. H. & Leake, B. E. (eds) *The Caledonides of the British Isles—Reviewed*, 3-18 (Geological Society, London, 1979).

Scale of body pattern adjusts to available cell number in amphibian embryos

Jonathan Cooke

National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

In many embryos, the removal of cells whose descendants would normally have formed entire parts of the body pattern is followed by apparently normal morphogenesis, which implies an ordered readjustment of the activities of the remaining cells before their potentialities become restricted. Special cell lineages cannot underlie the generation and regulation of pattern in such embryos^{1,2}. It is proposed instead that there must be some regulatory communication system in the developing embryo that normally ensures an appropriate spatial pattern of differentiation but which is also able to adjust to the removal, addition or transposition of material at a sufficiently early stage³⁻⁶. Precise models for such a mechanism have recently been suggested⁶⁻¹⁰, and have been tested experimentally¹⁰⁻¹⁴. I have performed surgical manipulations at pre-gastrula embryonic stages in two distantly related amphibian types, *Xenopus* and *Ambystoma*, and report here an assessment of the regulation achieved in terms of pattern proportions. The results are problematical for most current theories of pattern control.

Figure 1a depicts the commitment of cells to form four elements in the medio-lateral dimension of primary mesodermal pattern, during gastrulation and neurulation. Notochord is situated middorsally, and the pattern progresses laterally through somite and pro-nephros to lateral plate for the ventrally centred remainder of the cylindrical cell sheet. The pattern is probably set up in a dorso-ventral time sequence between the onset of gastrulation and the early tailbud stage¹⁵ (that is, 10–30 h, depending on species and temperature), during which about a doubling of cell number occurs along the dimension concerned². About 10^4 cells are involved, and pattern forms across some 1–2 mm (60–100 cells) from dorsal to ventral midlines.

Figure 1b describes the first experimental challenge to the regulatory system, the manufacture of artificially small (few-celled) blastulae that develop into early larval stages without compensatory change in mitotic schedule among the remaining cells. Fate mapping¹⁶ indicates that removed blastomeres would normally have contributed pattern extending from the ventral midline, sometimes to include the entire ventral pair of elements. Figure 1c shows transverse sections of normal and miniature sibling embryos at the late tailbud stage, at which the differentiating pattern was assayed. Embryos were scanned for cell number incorporated into each pattern part, between set positions in the antero-posterior axis of pattern, that is, the anterior end of the pro-nephros and the posterior gut opening as in Fig. 1d. Scans of about 25% of this cell population (using regular subsets of the serial transverse sections and an Abercrombie correction factor) revealed that morphologically small embryos contained 25–60% of the mesodermal cell

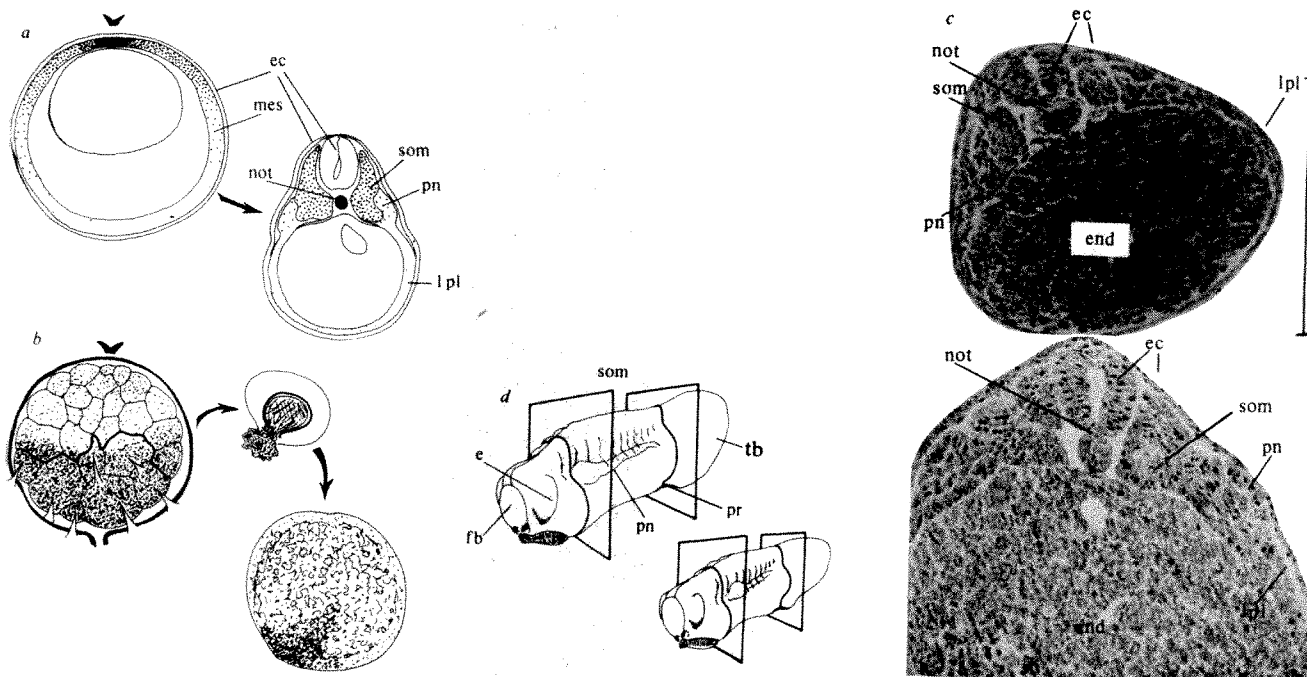


Fig. 1 a, The determination of pattern. Density of stippling in the middle mesodermal layer of the gastrula on the left, in transverse section, shows the classically recognized distribution of fates for the cells (in normal development), as they migrate forward (at right angles to the plane of the page) in the form of a coherent, roughly cylindrical, one cell-thick sheet. On the right, the same stippling represents the determined and anatomically visible zones of mesodermal structure at later (post-neurula) stages. Although these structures are solid due to piling up of cells towards the dorsal midline, their proportions in terms of cell numbers probably represent determinative events occurring in the earlier, essentially monolayered situation. Notochord (not) is a small (4%) rod-shaped mid-dorsal population, while paired somite (som) and mid-ventral lateral plate (lpl) zones each comprise ~40% of the cells. The remainder is devoted to the paired pro-nephric (pn) zones lateral to the somites. Determination seems to proceed in a dorso-ventral time sequence, starting with notochord, near the onset of gastrulation, but interactions that accompany grafting, or the co-culture of isolates, can alter development from more ventral regions until neurula stages. ec, ectoderm and neural structures induced from these; mes, mesoderm. b, Size reduction; ventral blastomeres were destroyed by pricking with a fine tungsten needle in an animal-to-vegetal sector, at morula or very large-celled blastula stages, with vitelline membrane and a thin jelly coat remaining on the embryo. Wax-coated dishes were used, containing 20% Niu-Twitty solution adjusted to pH 7.2 with dilute HCl. Dying material loses adhesivity for the healthy blastomeres and is squeezed out through an appropriately positioned slit in the coats, with rounding up and healing of the reduced blastula. Such blastulae were demembrated and kept with normal siblings in 10% Niu-Twitty on glass, until stages of fixation. The healed embryo, which reveals the loss of material in its size and reduced area of dark pigmentation, is shown to scale with the early gastrula section (in the appropriate plane) of a. ∇ marks mid-dorsal in a and b. c, Transverse section of small and normal sibling *Ambystoma* embryos at mid-trunk levels, as scanned in cell number sampling. Counts of nuclear density in the various mesodermal zones in this and the horizontal planes of section reveal that the adaptation of pattern to scale of the tissues is a genuine reflection of the participation of fewer, normal-sized cells in each structure. end, yolky endodermal mass. Scale bar, 1 mm. d, The sampling method. A control and an operated embryo are shown, with sectional levels that bounded the region of the axis used to assay medio-lateral pattern proportions. Sampling thus scanned the mesoderm in a uniform way in each embryo but excluded the head and heart region anteriorly and the tailbud posteriorly, where relatively small proportions of the whole population are involved in morphologically complex situations. Scale regulation definitely occurs in the longitudinal aspect of head and body pattern²⁴. e, eye rudiment; fb, forebrain region; pn, outline of pro-nephros; som, outline of somites; tb, tailbud; pr, protodaeum or original blastopore.

Table 1 Proportions in experimentally small embryos, in the patterns of original polarity in doubly organized embryos, and in control siblings

Species	Sibling set	% Of control mesodermal cell number	% Of control number of cells traversed in cross-section for small embryos, or % of host cross-sectional population involved for double dorsals	% Mesoderm cells devoted to:			Lateral plate
				Notochord	Somite	Pro-nephros	
<i>Ambystoma</i> small embryos	1 Control			3.7	44	10	42
	Exp.	51	69	4.5	44	12	39
	2 Control			4.2	39	11	46
	Exp.	{ 36 25	58 35	5.5 5.0	40 39	14 ND	40 56
	3 Control*			4.1	45	8.7	42
	Exp.	{ 41	59	3.7 4.8	38 39	11 11	47 45
	Double-dorsal embryos			3.8	50	10	36
	1 Control	{		4.1	46	11	39
	Exp.	{	74 56	4.2 4.6	43 48	9.4 12	43 35
	2 Control			6.0	51	8.7	34
<i>Xenopus</i> small embryos	1 Control*			3.9	37	10	49
	Exp.	{	64 59	4.3 4.1	43 36	12 9.5	41 50
	2 Control			5.0	36	15	43
	Exp.	{					
	3 Control*			4.2	43	6.8	46
	Exp.	{ 45	68	4.0	41	9.0	46
	2 Control*			4.0	37	11	48
	Exp.	{ 61 39	72 52	5.2 4.0	46 33	6.8 10	42 53
	3 Control			4.5	40	1.5	54
	Exp.	{ 30	61	4.0	44	11	41
	4 Control			4.0	39	8.0	49
	Exp.	{ 42	60	4.0	42	10	44
	Double-dorsal embryos			4.1	45	7.0	44
	1 Control			4.0	37	9.2	50
	Exp.	{	58	5.0	44	7.8	43
	2 Control			4.1	39	8.4	48
	Exp.	{	68 67	5.1 6.0	43 44	11 7.8	41 42
	3 Control			6.0	44	8.4	47
	Exp.	{	63 63 65 60	4.2 6.7 5.9 5.1	40 44 45 49	8.4 9.0 9.7 7.6	47 40 39 38

For the method of scanning patterns, see text and Figs 1a, d and 2b. ND, not differentiated.

* Mesoderm population within controls differed over about a 20% range; the mean being used to derive % values for experimentals within each set. Two significant figures are used throughout, notochord and pro-nephros being small structures, somite and lateral plate large.

number in synchronous control siblings, leading to patterns in which 30–70% of the normal numbers of cells were transected per section in the pattern dimension being considered. Structures form essentially in paired strips or zones between notochord and lateral plate, so that the results of pattern regulation can be expressed as percentages of cells devoted to each structure, relative to those in control siblings.

Cell numbers and the proportions devoted to each pattern element can be assessed in the same way after a second type of operation, the grafting of a dorsal blastoporal lip or 'organizer' region from a donor early gastrula to a site in the ventral marginal zone of a late blastula host^{13,17} (Fig. 2a). In these circumstances the embryo differentiates mesodermal patterns centred round two typical dorsal midline structures, but with full cellular continuity at their lateral interface or plane of mirror symmetry within the original cylinder of the mesodermal mantle. This is because the grafted organizer, while adhering to its own presumptive fate in development as the dorsal midline, has also interacted during gastrulation with the host mesodermal cell population to organize a second axial pattern, embracing ~40% of that population and centred on itself. Cell counts as in Fig. 1d show that the total population within these mirror-symmetric mesoderms is indistinguishable from that in normal siblings at early differentiated stages as studied here (see also ref. 2). Thus, as after the previously described operation, mitotic schedule has been unaltered during cell interactions that adjust pattern.

Table 1 gives pattern proportions in embryos of every set so far examined. Because implanted organizers have only belated opportunity to interact with surrounding material, compared

with the host's original one, most models for the regulatory mechanism can explain how these may involve less than half the host tissue in a pattern of new polarity, as is usually observed. Attention is therefore confined in Table 1 to proportions in the single patterns of miniature embryos and to those in the patterns of original polarity, around the host dorsal midline, in the doubly organized ones. The behaviour within these latter patterns has the most distinctive relevance in assessing models (see discussion below). However, note that when cell numbers are pooled from both patterns for each structure, embryos with two dorsal midlines have only a slight under-representation of 'ventral' parts compared with normal, whereas an obliteration of at least the lateral plate territory might have been expected due to involvement of its tissue in new dorsal parts.

The following points emerge from the data: (1) There is appreciable variation in proportion even among normal sibling embryos, perhaps exceeding that estimated during formation of the proximo-distal pattern of the bird limb¹⁸. Nevertheless, as populations of observations, proportions within the homologous patterns at equivalent stages are remarkably similar within these two distantly related species. (2) Regulation is highly effective in assigning appropriate numbers of the available cells to the four pattern parts. Pro-nephros is relatively the most variable in proportions, in normal development and particularly in very small patterns where it may hardly exist. An apparent tendency for lateral plate to be over-represented in cases where pro-nephros is under-represented may imply that the latter is determined after somite, in tissue which will otherwise become lateral plate. If we therefore estimate pattern as three elements only, that is, notochord, somite and pro-nephros plus lateral plate,

normal and experimentally small embryos become highly comparable populations in terms of proportions. Only notochord, probably determined soon after onset of gastrulation, tends to be rather over-represented where the scale is small. Regulation has been strikingly successful, as obliteration of lateral plate and often of pro-nephros would otherwise have resulted because of the particular cells deleted. A quantitative estimate of the degree of regulation after these experimental challenges is beyond the scope of this study due to the level of variation in the outcome of undisturbed pattern formation. (3) A new organizer, situated ventrally, exerts an active influence in scaling down the pattern formed, with normal polarity around the original dorsal midline, so that both contrapolar versions of the pattern may be almost complete despite sharing a normal-sized tissue sheet during their formation. This requires, as observed, that the pattern centred round the host organizer (as well as that around the new one) has fewer cells in its dorsal parts than do normal versions.

Recent models of control of pattern in early development have arisen mainly from the notion of positional information⁶. This postulates a cell variable or signal, describable as a quantity (for example, morphogen substance concentration), and distributed in a monotonically graded profile between set 'bound-

dary' values across the available tissue for each pattern dimension. Cells are held to interpret this gradient by adopting developmental pathways, or determined states corresponding to particular pattern elements, in response to experience of particular absolute ranges of the signal variable. Thus stated, the idea is a formalism independent of hypotheses as to underlying mechanisms; in Fig. 3 the formalism is used to represent the normal medio-lateral pattern (a) and those actually seen after the two experimental challenges described here (version 3 of b and c). The gradient concept agrees well with the order and polarity shown by the sequences of parts in patterns, and for the

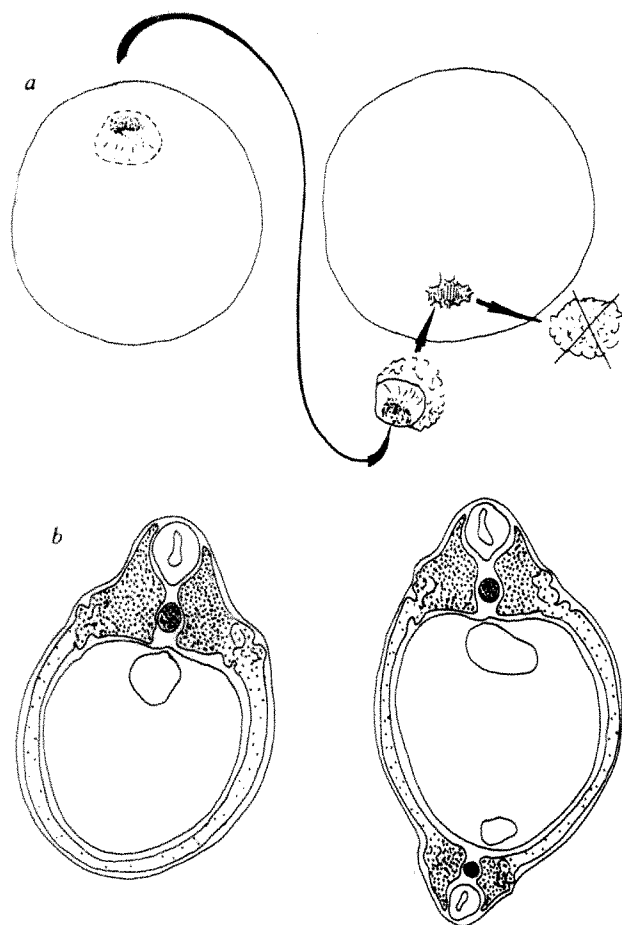


Fig. 2 a, Early gastrula donor and late blastula host viewed from vegetal aspect, presumptive dorsal midline uppermost, to show the implantation of the cell group from the dorsal blastoporal lip into a ventral marginal position at a prepared site. In the control sham operation, which is followed by normal development, the donor site is simply excised and replaced with another, similar cell group. b, Schematic transverse sections (as in Fig. 1a) through a normal embryo, and one having developed a second dorsal axis of pattern through implantation of a dorsal lip, as in a. The axial level is nearer the anterior of the two boundaries indicated in Fig. 1d. Both experimental patterns, though sharing one cell population normal for the embryo's age, have approximately normal proportions. Results in Table 1 include only proportions for the pattern retaining its original polarity in these embryos, that is, the one orientated normally in the right-hand diagram. The effect of the new organizer on the scale of structures in the pattern around the old one is the most troublesome feature for models that would explain the results on the basis of positional information.

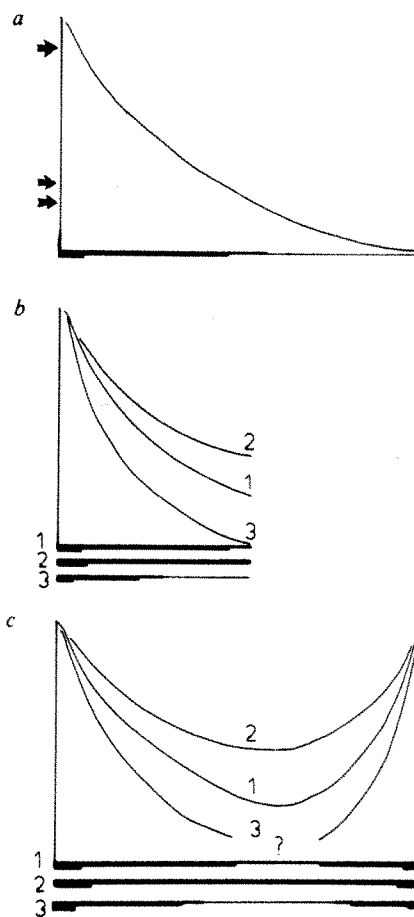


Fig. 3 a, A gradient profile which should be taken to represent formally, in scale and proportions, the normal medio-lateral pattern within the mesoderm at determination. The four thicknesses along the baseline represent the extents of zones devoted to the four pattern parts, in terms of cell number within the originally monolayered mesodermal cylinder. Arrows on the ordinate mark threshold values for a positional variable that would set the boundaries between parts on the positional information paradigm. The monotonic gradient is depicted as concave nonlinear because most proposed mechanisms would give such a profile for, say, morphogen concentration after diffusion from a set (organizer) source. The present experimental system offers no evidence for any dynamic second organizing boundary corresponding to a local 'sink'. b, The experimental challenge of small embryonic size. Profile versions 1 and 2 represent expected results on some simple diffusion models (see text), while version 3 shows the observed result, and therefore the gradient profile that any true positional information will have assumed in order to produce it. c, The experimental challenge of double organization in a normal-sized field. Profile and baseline versions 1 and 2 again represent the range of results (from non-interaction, to mutual 'flooding out' of gradient values) expected from the simplest model mechanisms. Version 3 represents the actual behaviour and thus the necessary gradient profile if positional information is to be used. A true positional gradient for this vertebrate pattern dimension would thus have to (1) convey significant local information across much of its extent, and therefore be not extremely nonlinear (note that pro-nephros is both correctly positioned and significantly size-regulated across a range of scales, and is near the midpoint of pattern) and (2) be able to restore its lowest values to cells normally experiencing middle values, when the normal low cells were removed or replaced by a second upper boundary. These are demanding requirements, and alternative theories of pattern control might be more parsimonious, and potentially testable.

phenomenon of the organizer (that is, a pre-determined 'upper boundary' cell group controlling each pattern). But concrete models for positional information mechanisms must be able to explain, or produce during computer simulations, gradient landscapes whose scales and profiles would give the empirically seen results when interpreted as in the formalism.

Models where the organizer is a source at a pre-set level for a diffusing morphogen, and cells elsewhere are sinks, are inadequate. In such schemes, each set of parameters for diffusion, synthesis and destruction of morphogen, implies a gradient or a set time sequence of gradients that occupy particular spaces when embracing a particular concentration range. They would thus lead to profiles between versions 1 and 2 of Fig. 3b and c, corresponding to lack of regulative replacement of ventral pattern, or even to the expansion of dorsal parts due to 'flooding out' of ventral positional values. Models which use reaction-diffusion mechanisms^{8,19,20} are better, as they can produce certain size-regulated gradient profiles, and steepen gradient slopes in systems that have become bipolar in a way that would correspond with present results. Their capacities are stretched to the limits of plausibility, however, by the shape of gradient that would be required here (see Fig. 3 legend) and by the lack of evidence that any region develops 'ventral' organizer properties complementary to those of the dorsal lip. Models where each organizer acts initially to polarize morphogen transport in surrounding tissue, and thus produce its gradient, lead most easily to the necessary gradient profiles. However, there is no indication of structure or physiology that could underlie massive 'uphill' substance transport across embryos.

Alternatively, we can question the hypothesis that cells form the pattern by interpreting a positional gradient alone. Diffusible signal substances made within and specific to each developing pattern part could be involved, acting as size/proportion sensors for the system because the rate at which they are made depends on how many cells are proceeding towards a particular determination, whereas the rate at which they are destroyed depends on total tissue present. These might control the sizes of elements by feedback inhibition that switches cells towards each successive state in a series. The role of each organizer would then be simply to ensure that onset of cell determination *per se*, leading to production of the succession of specific inhibitors, occurs as a coherent (medial to lateral) wavefront across tissue. Demands for sophisticated regulation in such a gradient system would be avoided. This type of model, further described elsewhere^{21,22} makes predictions which are testable and contrast with those of positional information theory²¹, as well as accounting for interaction between organizers in scaling pattern parts (see Fig. 3c); see ref. 23 for discussion of formally similar models.

I thank June Colville for histology, and the MRC for supporting this work.

Received 29 January; accepted 11 March 1981.

1. Dan-Sohkawa, M. & Sato, N. *J. Embryol. exp. Morph.* **46**, 171-185 (1978).
2. Cooke, J. J. *Embryol. exp. Morph.* **51**, 165-182 (1979).
3. Driesch, H. *The Science and Philosophy of the Organism*, 2nd edn (Black, London, 1929).
4. Weiss, P. Q. *Rev. Biol.* **15**, 177-198 (1950).
5. Waddington, C. H. *J. exp. Biol.* **15**, 377-381 (1938).
6. Wolpert, L. *J. theor. Biol.* **25**, 1-48 (1969).
7. Goodwin, B. C. & Cohen, M. H. *J. theor. Biol.* **25**, 49-107 (1969).
8. Gierer, A. & Meinhardt, H. *Kybernetik* **12**, 30-39 (1972).
9. Meinhardt, H. *Rev. physiol. biochem. Pharmac.* **80**, 47-104 (1978).
10. Tickle, C., Summerbell, D. & Wolpert, L. *Nature* **254**, 199-202 (1975).
11. Wolpert, L., Clarke, M. R. B. & Hornbruch, A. *Nature* **239**, 101-105 (1972).
12. Summerbell, D. *J. Embryol. exp. Morph.* **50**, 217-233 (1979).
13. Cooke, J. J. *Embryol. exp. Morph.* **28**, 13-26, 27-46, 47-56 (1972); **30**, 283-300 (1973).
14. Herth, W. & Sander, K. *Arch. EntwMech Org.* **172**, 1-27 (1973).
15. Forman, D. & Slack, J. M. W. *Nature* **286**, 492-494 (1980).
16. Keller, R. E. *Dev. Biol.* **51**, 118-137 (1976).
17. Spemann, H. & Mangold, H. *Arch. microsc. Anat. EntwMech.* **100**, 599-638 (1924).
18. Summerbell, D. & Wolpert, L. *Nature* **244**, 228-229 (1973).
19. Sibatani, A. *J. theor. Biol.* **75**, 149-166 (1978).
20. Papageorgiou, S. *Biophys. Chem.* **11**, 183-198 (1980).
21. Cooke, J. *Am. Zool.* (in the press).
22. Rose, S. M. *Growth* **31**, 149-164.
23. Meinhardt, H. & Gierer, A. *J. theor. Biol.* **85**, 429-450 (1980).
24. Cooke, J. *Nature* **254**, 196-199 (1975).

Sequential functions of the bithorax complex of *Drosophila*

Gines Morata & Stephen Kerridge

Centro de Biología Molecular, Universidad Autónoma de Madrid, Canto Blanco, Madrid 34, Spain

The bithorax genes are a group of homoeotic genes of *Drosophila* whose function is related to the control of segment development. They form a gene complex in which at least seven different loci have been identified¹⁻³. Mutations at each of these loci produce a specific homoeotic transformation whereby a segment is transformed either in part or completely into another³⁻⁵. In embryos completely deficient for the bithorax genes, all the thoracic and abdominal segments resemble mesothoracic segments³. This observation, together with the segmental specificity of the mutant phenotypes, led Lewis³ to suggest a model of genetic control in which the type of development in each segment is specified by the activation of a fraction of the total number of bithorax genes. Thus, lack of activity of the bithorax genes results in mesothoracic development, activation of the *bithorax* (*bx*⁺) and *postbithorax* (*pbx*⁺) genes produces metathorax (anterior and posterior respectively), and activation of these two genes with *bithoraxoid* (*bxd*⁺) produces the first abdominal segment and so on. Although this model explains the genetic and developmental data, it leaves unexplained formation of the prothoracic segment. We now describe two unexpected results: deficiencies for bithorax genes can lead to segments being transformed into prothorax, and there is a temporal sequence in the function of the bithorax system.

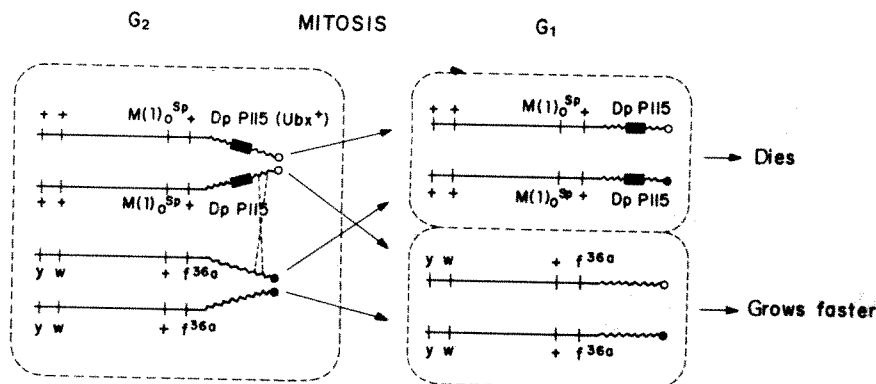
We have studied the development of mutants in the locus *Ultrabithorax* which behave as deficiencies for some of the bithorax genes^{1,2,6}. *Trans* combinations of *Ubx* alleles with any of the recessive bithorax pseudoalleles produce the phenotype of the recessive allele in question. As the *Ubx* alleles, like all bithorax deficiencies, are zygotic lethals, they cannot be studied in homozygous adults. Thus, description of the bithorax deletions has so far been restricted to the phenotype before death of embryos or first instar larvae³. Here we have used an alternative approach, whereby we induced clones of cells mutant for *Ubx* in fully viable heterozygous embryos or larvae and examined the phenotypes in the cuticle of adult flies, whose segment patterns are better known and easier to distinguish than those of the embryo or larva. Most of the zygotic lethals are cell viable in the adult cuticle⁷, and mutant clones can be recognized because they are labelled with cuticular markers (see Fig. 1). This method has the advantage that mutant clones can be produced at any desired developmental stage. In most experiments we used the *Minute* technique⁸ to obtain very large clones in which the phenotype can be clearly defined.

We chose the mutation *Ubx*¹³⁰ for most of the experiments because it produces very strong mutant phenotypes over *bx*, *pbx* and *bxd*. To avoid the possibility of effects of other mutants in the *Ubx*¹³⁰ chromosome, the flies used in the clonal analyses carried *Ubx*¹³⁰ over the deletion *Df(3R)P9*. This lethal genotype was rescued by a duplication in the first chromosome that carries all the bithorax genes in wild-type condition (see Fig. 1 legend for technical details).

Clones were induced at 4 ± 2, 7 ± 1, 17 ± 1, 25 ± 1 and 58 ± 8 h after egg laying (AEL). Those induced at 4 ± 2 h AEL will be considered as blastoderm clones, and the rest post-blastoderm.

We examined blastoderm clones all over the body. In dorsal structures they differentiated normally in all segments except the metathoracic and first abdominal ones. In agreement with previous results^{9,10}, clones in the haltere showed a transformation into wing, because *Ubx*¹³⁰/*Df(3R)P9* is deficient for bithorax and postbithorax activities. For reasons discussed

Fig. 1 Details of mitotic recombination to produce clones of cells defective in bithorax activity in $M(1)o^{Sp}Dp(3;1)P115/y w f^{36a}$; $In(3LR)Ubx^{130}$, $Ubx^{130e^s}/Df(3R)P9$ flies. The lethality and mutant phenotype of the combination $Ubx^{130}/Df(3R)P9$ in the third chromosome is rescued by $Dp P115$ which carries wild-type alleles of the bithorax genes. The diagram shows a mitotic recombination event in the first chromosome (the genetic constitution of the third chromosome remains unchanged). Recombination proximal to $Dp P115$ results in the loss of the wild-type bithorax genes in one of the daughter cells. This cell is simultaneously marked with $y w$ and f^{36a} (see ref. 12) and also loses the Minute condition, which confers a growth advantage to the marked cell. The other daughter cell dies due to the presence of $M(1)o^{Sp}$ homozygously⁸. Note that mitotic recombination distal to $Dp P115$ and proximal to f^{36a} produces clones marked with $y w f^{36a}$ that remain wild type for bithorax. We estimate these are 20–25% of the total number of $y w f^{36a}$ clones (see Tables 1 and 2).



elsewhere⁴, thoracic cells cannot differentiate if they are in an abdominal system. Thus, the absence of mutant clones in the first abdominal segment indicates *bxd* transformation. Therefore, the phenotype of the $Ubx^{130}/Df(3R)P9$ clones in the dorsal parts is the one expected for the triple deficiency for *bx*, *pbx* and *bxd*.

However, clones found in the ventral thoracic structures (legs), where bristle patterns can be distinguished as pro-, meso- or metathoracic⁹, showed one important difference compared with those found dorsally—although they all respected the antero-posterior boundary line, the posterior clones of the second (IIP, mesothoracic) and third (IIIP, metathoracic) legs showed a transformation into posterior first (IP, prothoracic) leg (Fig. 2). Clones in the first leg remained untransformed (Table 1). This is the first evidence of transformation into prothorax by a mutation in the bithorax system. Note also that two different segments, mesothorax and metathorax, show the same transformation. The clones in anterior compartments behaved differently—in the first and second legs they remained untransformed but in the third leg they showed a transformation into second. This is the phenotype of mutations in the *bx* locus.

Postblastoderm clones were induced at 7 ± 1 , 17 ± 1 , 25 ± 1 and 58 ± 8 h AEL (Table 1), the last time corresponding to the middle of the second larval period. Clones were found in all body segments and their phenotypes were as described for the blastoderm clones except that at 17 h AEL and at later times IIP and IIIP clones were no longer transformed into IP. Instead, IIP clones remained as IIP and those in IIIP were transformed into IIP; that is, they showed *pbx* phenotype. At 7 ± 1 h AEL some clones showed prothoracic and others a mesothoracic pattern. Thus, the transition period between prothoracic and mesothoracic transformation is at about 7 h of development. These results have been confirmed with nine new *Ubx* alleles induced in our laboratory¹¹ and also with Ubx^1 , which, unlike Ubx^{130} , is a point mutant not associated with any chromosomal rearrangement^{1,2,12}. In this latter case we found four blastoderm clones in IIIP and two in IIP: all transformed into IP. We have also studied the legs of a sample of 32 gynandromorphs (lent by M. P. Capdevila and F. Miñana) in which the male part was mutant for Ubx^1 . The phenotype of the gynandromorph clones was the same as the blastoderm clones.

Thus clones of the same genotype, occupying the same position in the adult structure, show different mutant phenotype depending on the time of clone induction. This can be best explained if the clones are deficient for several functions which are required in a temporal order. It also suggests the existence of an unsuspected gene in the bithorax system, required for the normal development of the mesothoracic and metathoracic posterior legs.

To test whether this gene is different from *pbx* which acts in the same position, we tested whether partial duplications of the

bithorax complex can cover the posterior prothoracic transformations but not the *pbx* transformation. We used $Tp(3)bxd^{100}$ (see ref. 12), a euploid chromosome where some but not all the bithorax genes (*bx*, *Cbx* and *Ubx* but not *bxd* and *pbx*) are transposed to the left arm of the third chromosome. It has a breakpoint between polytene bands 89E2–3 (ref. 12) (within the bithorax region, which extends cytologically from 89E1 to E4) and is mutant for *pbx*. By meiotic recombination we can replace the left arm of this chromosome by a normal one generating a deletion for all transposed genes. This chromosome, $Df(3R)bxd^{100}$, is also mutant for *pbx* and

Table 1 Homoeotic transformations shown by $y w f^{36a}$; $Ubx^{130}/Df(3R)P9$ clones found in the posterior compartments of the legs of $M(1)o^{Sp}Dp(3;1)P115/y w f^{36a}$; $In(3LR)Ubx^{130}$, $Ubx^{130e^s}/Df(3R)P9$ females after irradiation (500 rad) at different times in development

Age at irradiation (h AEL)	No. sides	Phenotype of the clones found in		
		IP leg	IIP leg	IIIP leg
4 ± 2	1,332	IP(22)	IP(13)	IP(14)
7 ± 1	424	IP(7)	IIP(2)	IIP(3)
			IIP(9)	IP(1)
				IIP(2)
				IIIP(1)
17 ± 1	2,030	IP(36)	IIP(36)	IIP(35)
25 ± 1				
58 ± 8				

Data from irradiation at 17 ± 1 , 25 ± 1 and 58 ± 8 h AEL are pooled as there was no difference between them. The actual number of clones observed is shown in parentheses. In IIP and IIIP where homoeotic transformation is seen there is always a fraction of clones that remain untransformed. These are probably the result of mitotic recombination distal to $Dp(3;1)P115$ and proximal to *forked* (see Fig. 1). Dorsal clones induced in this experiment were untransformed except in the metathoracic and first abdominal segments. Clones in the haltere were transformed into wing, anterior clones could completely reproduce the anterior wing compartment (*bx* transformation) and posterior clones could make a posterior wing compartment (*pbx* transformation). In no case did a clone cross the antero-posterior compartment boundary, either in the metathoracic where they produce a transformation or in mesothoracic segment where they have no homoeotic effect. No mutant clones were found in the first abdominal segment. This result is similar to that described⁴ for clones homozygous for Ubx^1 and for bxd^1 induced during the larval period, and probably indicates the transformation of first abdominal histoblast cells into metathoracic or mesothoracic cells. As a control we looked for clones in sib flies of the genotype $Dp(3;1)P115 M(1)o^{Sp}/y w f^{36a}$; $Dp(3;3)P5/In(3LR)Ubx^{130}$, Ubx^{130e^s} where a mitotic event of the type described in Fig. 1 produces $y w f^{36a}$ clones that remain wild type for the bithorax complex. None of the clones found, either in legs (72 in total) or elsewhere in the body, produced homoeotic transformation. To test the remote possibility that transformation into prothorax could be caused by the excess proliferation of M^+ clones which could interact in some unexpected manner with surrounding slowly proliferating Minute cells, we irradiated at 4 ± 2 h AEL embryos of the genotype $Dp(3;1)P115/y w f^{36a}$; $Df(3R)P9/In(3LR)Ubx^{130}$, Ubx^{130e^s} . In this case the $y w f^{36a}$ clones proliferate at the same rate as surrounding cells. Out of five clones in IIP leg, three were clearly transformed into IP and the three clones in IIIP were transformed into IP. Thus the prothoracic transformation does not result from a difference in proliferation rate.

generates an equally strong *pbx* transformation to that of *Tp(3)bx¹⁰⁰/pbx*. We expected the wild-type gene responsible for the prothoracic transformation to be missing in *Df(3R)bx¹⁰⁰* but present in *Tp(3R)bx¹⁰⁰*, in which case blastoderm clones of the genotype *Df(3R)bx¹⁰⁰/Ubx¹³⁰* would show prothoracic transformation whereas blastoderm *Tp(3R)bx¹⁰⁰/Ubx¹³⁰* clones would not. In both genotypes post-blastoderm clones should show a *pbx* phenotype.

The experiments were conducted by the same method as those described for *Df(3R)P9*, substituting *Tp(3)bx¹⁰⁰* or *Df(3R)bx¹⁰⁰* for *Df(3R)P9*. The results are shown in Table 2 and they are exactly as expected: the prothoracic transformation is absent in *Tp(3;3)bx¹⁰⁰/Ubx¹³⁰* clones but present in

Table 2 Type of homoeotic transformation shown in the posterior compartments of the legs by blastoderm clones of the genotypes indicated

Genotype	No. sides	Phenotype of blastoderm clones found in		
		IP leg	IIP leg	IIIP leg
<i>Df(3R)bx¹⁰⁰/Ubx¹³⁰</i>	910	IP(7)	IP(6) IIP(1)	IP(10) IIIP(1)
<i>Tp(3;3)bx¹⁰⁰/Ubx¹³⁰</i>	976	IP(7)	IIP(8)	IIP(6) IIIP(1)

Clones were induced after irradiation (500 rad) at 4 ± 2 h AEL in female embryos of genotype *y w f^{36a}/M(1)0^{5p}Dp(3;1)P115; In(3LR)Ubx¹³⁰, Ubx¹³⁰e⁺/Df(3R)bx¹⁰⁰* in the first case and *y w f^{36a}/M(1)0^{5p}Dp(3;1)P115; In(3LR)Ubx¹³⁰, Ubx¹³⁰e⁺/Tp(3;3)bx¹⁰⁰* in the second. Both genotypes are phenotypically wild type for bithorax. The number of clones found is shown in parentheses. Some clones in legs IIP and IIIP did not show transformation probably for the reason mentioned in Fig. 1 and Table 1. Post-blastoderm clones of both genotypes showed mesothoracic transformation.

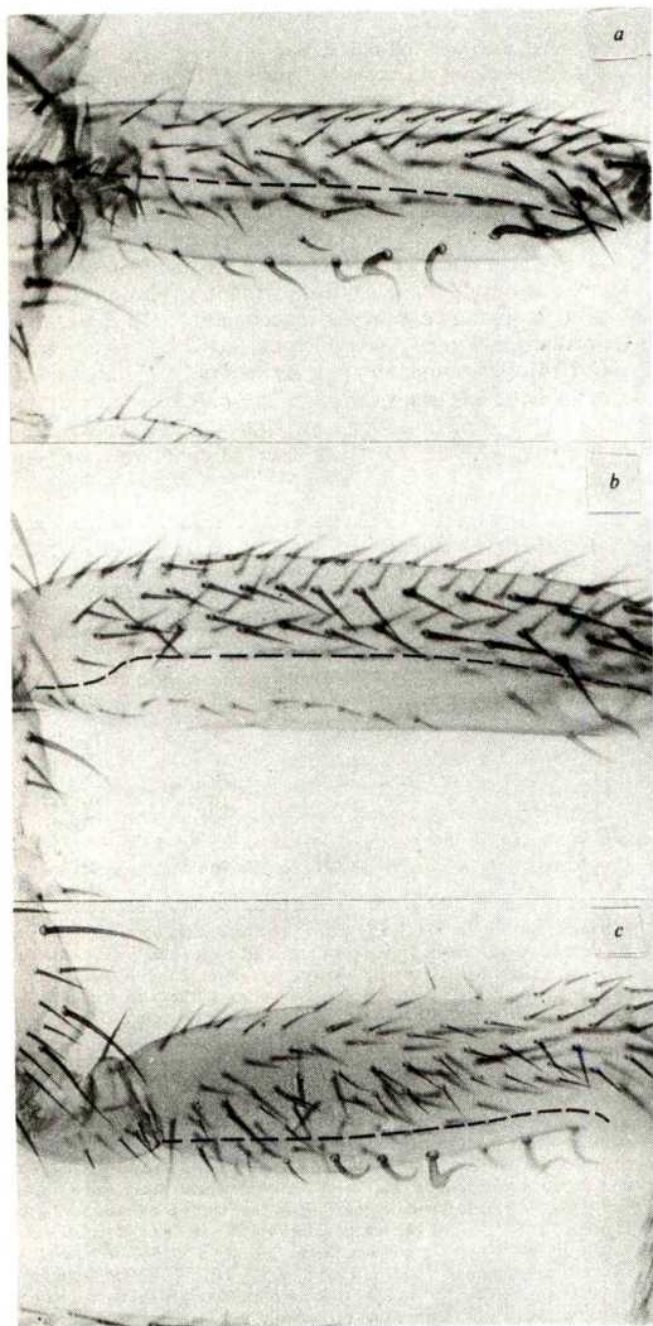


Fig. 2 Homoeotic transformation shown by *y w f^{36a}; Ubx¹³⁰/Df(3R)P9* clones (below the dotted line) in the posterior compartment of the second leg. *a*, Blastoderm clone showing transformation into posterior first leg; *b*, post-blastoderm clone not showing transformation. Note the difference in the bristle pattern of the two clones. *c* Shows a blastoderm clone in the posterior first leg for comparison.

Df(3R)bx¹⁰⁰/Ubx¹³⁰. Therefore, the gene responsible for the prothoracic transformation is present in the transposed fragment and is not associated with *pbx*. It was previously thought that lack of bithorax function results only in mesothoracic development³. However, these results show the existence of a gene of the bithorax system that, if mutant, transforms the posterior compartments of both mesothoracic and metathoracic legs into prothorax. The gene is located to the left of *pbx* and included in the fragment of the bithorax complex missing in *Df(3R)bx¹⁰⁰*. Its function is necessary for the development of the mesothoracic and metathoracic segments and its realm of action is defined by the antero-posterior compartment boundary. Like the genes *bx*, *pbx* and *bx^d*, this one also is inactivated by *Ubx* alleles. Following Lewis' nomenclature we propose to call this gene *postprothorax* (*ppx*). Its period of function differs from that of *bx* and *pbx*; *Ubx¹³⁰/Df(3R)P9* clones show prothoracic transformation only if induced before 7 h AEL. Clones of the same genotype induced at 17 h AEL develop as mesothorax. Thus, *ppx* function is needed in early development but is dispensable thereafter, unlike *bx* and *pbx* whose functions are needed until the end of the larval period⁴.

The two sets of genes seem to act sequentially: in the posterior metathoracic leg there is an early requirement for the *ppx* gene activity, which blocks prothoracic development and directs the cells to the mesothoracic pathway. A few hours later *pbx*⁺ gene function is required, preventing mesothoracic and allowing metathoracic development. Thus the final developmental route taken by the posterior metathoracic leg is specified by the combined sequential function of *ppx* and *pbx*. This suggests a temporal mechanism in the genetic process of thoracic and abdominal determination. Each segment would undergo a succession of determinative steps, each genetically implemented by the activation or inactivation of a particular homoeotic gene. The final development of each segment would be specified by a unique combination of active genes.

The *ppx* transformation affects only ventral structures (legs) whereas *bbx* (and *bx*) transforms the dorsal (halteres) as well. We cannot explain this difference but a common feature of all the homoeotic mutants producing transformation into prothorax is that this transformation is shown exclusively in the legs (for example *Polycomb* (*Pc*)¹³, *Extra sex comb* (*Scx*)¹³ and *Multiple sex comb*¹⁴).

Although our results only show evidence for a posterior prothoracic transformation, the existence of mutations like *bx* and *pbx*, that produce homologous mesothoracic transformations in anterior and posterior compartments respectively, suggests the existence of an 'anterior' prothoracic gene which, when defective, would transform the anterior meso- and metathorax into anterior prothorax. This hypothetical gene need not be in the bithorax complex.

The existence of prothoracic transformations produced by bithorax mutants apparently contradicts with results of Lewis³. He finds that embryos homozygous for *Df(3R)P9* show a transformation of all segments into mesothorax. Our explanation for this discrepancy is that ventral landmarks used by Lewis

(denticle belts) are all anterior and the prothoracic transformation could only be seen in the posterior compartments.

We thank Jim Kennison, Peter Lawrence, Dan Lindsley, Pedro Ripoll and Lenny Robbins for comments, Rosa González for technical help, and Victoria Lafita and Ernesto Sánchez-Herrero for typing the manuscript. S.K. is supported by an EMBO fellowship.

Received 7 November 1980; accepted 2 March 1981.

- Lewis, E. B. *Am. Zool.* **3**, 33–56 (1963).
- Lewis, E. B. in *The Role of Chromosomes in Development* (ed. Locke, M.) 231–252 (Academic, New York, 1964).
- Lewis, E. B. *Nature* **276**, 565–570 (1978).
- Morata, G. & García-Bellido, A. *Wilhelm Roux's Arch. dev. Biol.* **179**, 125–143 (1976).
- García-Bellido, A., Ripoll, P. & Morata, G. *Dev. Biol.* **48**, 132–147 (1976).
- Morata, G. & Kerridge, S. in *Development and Neurobiology of Drosophila* (eds Sidiqui, O., Babu, P., Hall, C. & Hall, J.) 141–154 (Plenum, New York, 1980).
- Ripoll, P. *Genetics* **86**, 357–376 (1977).
- Morata, G. & Ripoll, P. *Dev. Biol.* **42**, 211–221 (1975).
- Steiner, E. *Wilhelm Roux's Arch. dev. Biol.* **180**, 31–46 (1976).
- Lawrence, P. A. & Morata, G. *Dev. Biol.* **56**, 40–51 (1977).
- Kerridge, S. & Morata, G. (in preparation).
- Lindsley, D. L. & Grell, E. H. *Genetic Variations of Drosophila melanogaster* (Carnegie Institute, Washington, 1958).
- Hannah-Alava, A. *Genetics* **43**, 878–906 (1958).
- Tokunaga, C. & Stern, C. *Dev. Biol.* **11**, 50–81 (1965).

Immobilization of concanavalin A receptors during differentiation of neuroblastoma cells

Mark C. Fishman*, Paul R. Dragsten†‡ & Ilan Spector§

*Laboratory of Developmental Neurobiology, NICHD, National Institutes of Health, Bethesda, Maryland 20014, USA

†Laboratory of Theoretical Biology, NCI, National Institutes of Health, Bethesda, Maryland 20014, USA

§Laboratory of Biochemical Genetics, NHLBI, National Institutes of Health, Bethesda, Maryland 20014, USA

Neuroblastoma cells serve as a useful model of neuronal development because compounds such as dimethyl sulphoxide (DMSO) and dibutyryl cyclic AMP cause them to undergo a process of controlled differentiation in tissue culture, during which they can extend long processes, develop characteristic excitability mechanisms, synthesize neurotransmitters and form synapses^{1–3}. We have used the technique of fluorescence photobleaching recovery to study the lateral mobility of cell-surface constituents during the differentiation of neuroblastoma clone N1E-115 cells. The concanavalin A (Con A) binding sites appear as discrete patches distributed over the entire cell surface and exhibit lateral mobility in undifferentiated cells comparable with that of surface glycoproteins of other cells. After induction of differentiation, however, the vast majority of Con A binding sites become immobilized, and we present data which suggest that the mechanism of this immobilization may involve linkage to the internal actin network.

Cells of the adrenergic clone N1E-115 were grown and passaged in Dulbecco's modified Eagle's medium supplemented with 2.5% fetal bovine serum, in an atmosphere of 10% CO₂, 90% air, and replated in 35-mm plastic tissue culture dishes either with ('differentiated') or without (control) 2% DMSO. Cells were washed twice with Hank's balanced salt solution before labelling with the lectins succinyl Con A or wheat-germ agglutinin (WGA), which had been conjugated with tetramethylrhodamine isothiocyanate (E-Y Laboratories), or with the lipid analogue 3,3'-dioctadecylindocarbocyanine iodide (diI). Observation was with a Zeiss Universal microscope equipped with a $\times 63$ 1.25 n.a. water immersion lens and epifluorescence illumination.

Lateral mobility was evaluated quantitatively by the fluorescence photobleaching recovery technique^{4,5}, in which a laser beam is used as the excitation source and is sharply focused on a small area (1 μm^2) of the membrane of a labelled cell. A brief intense pulse then irreversibly bleaches 40–70% of the fluorophore within that region. Monitoring the recovery of fluorescence, as unbleached molecules subsequently diffuse into the bleached region, allows calculation of the relative amounts of mobile and immobile fluorophore and the diffusion coefficient of the mobile fraction.

Con A binding sites were distributed in patches over the entire cell surface, including soma, processes and growing tips. This clustering was apparent even if the cells were prefixed with 3.7% formaldehyde, or labelled at 4 or 37 °C. No bulk movement or capping of receptors was observed. Fluorescence intensity was dramatically reduced if labelled cells were subsequently

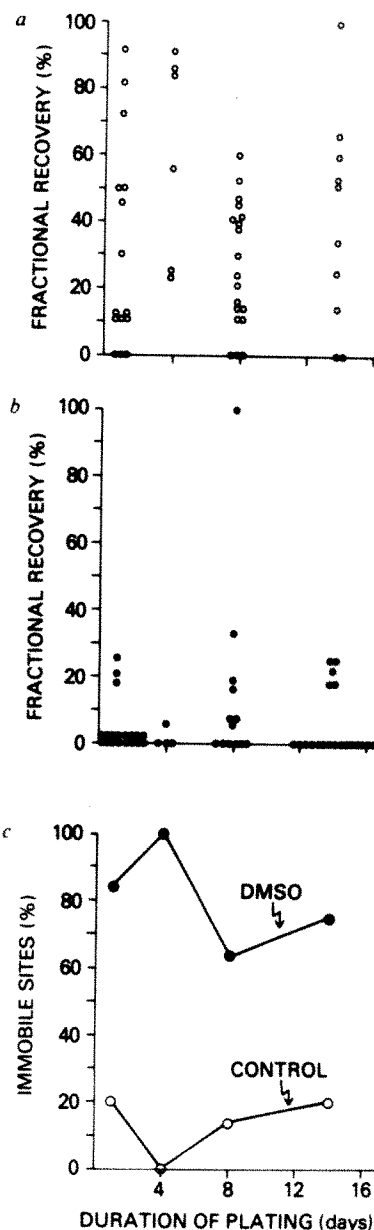


Fig. 1 Recovery of fluorescence after photobleaching. Cells were labelled with 12.5 $\mu\text{g ml}^{-1}$ rhodaminated succinyl Con A at room temperature for 20 min, and then washed twice with the balanced salt solution. Each point represents the per cent recovery of fluorescence after bleaching a site on cells incubated in control conditions (a) or after treatment with 2% DMSO (b) for the number of days indicated. c, The percentage of measurements showing immobilized Con A receptors (<10% recovery of fluorescence after photobleaching) when exposed to control (○) and differentiating (●) conditions for the number of days indicated.

‡ Present address: The Proctor and Gamble Company, Research and Development Department, PO Box 39175, Cincinnati, Ohio 45247, USA.

Table 1 Glycoprotein and lipid probe mobility on neuroblastoma cells

Probe	Treatment	Immobile sites (%)	$D \times 10^{10}$ ($\text{cm}^2 \text{s}^{-1}$)	Mean fractional recovery (%)
Succinyl Con A	Untreated (51)	18	1.3 ± 1.0	36 ± 28
	DMSO (all sites) (56)	79	—	7 ± 15
	Soma (31)	85	—	8 ± 20
	Hillock (13)	69	—	6 ± 8
	Process (12)	83	—	4 ± 8
	Dibutyl cyclic AMP (7)	71	—	5 ± 10
	Cytochalasin B DMSO (13)	23	1.6 ± 0.8	23 ± 18
	No DMSO (10)	10	1.0 ± 0.7	28 ± 20
	Colcemid DMSO (8)	75	—	10 ± 16
	No DMSO (10)	30	1.4 ± 1.0	29 ± 21
WGA	Untreated (12)	50	—	12 ± 12
	DMSO (14)	64	—	12 ± 13
	Cytochalasin B DMSO (14)	50	—	14 ± 12
diI	Untreated (10)	0	101 ± 20	—
	DMSO (9)	0	122 ± 50	—

Cells were labelled with $12.5 \mu\text{g ml}^{-1}$ succinyl Con A (10–20 min), $12.5 \mu\text{g ml}^{-1}$ WGA (20 min) or $10 \mu\text{g ml}^{-1}$ diI (3 min) and washed twice with the balanced salt solution. Treatment with DMSO was for 2–14 days, and with dibutyl cyclic AMP for 7 days. Cytochalasin B was used at a concentration of $10 \mu\text{g ml}^{-1}$ during the preincubation (30 min) and left in the dish during measurements. Colcemid was used at a concentration of $0.4 \mu\text{g ml}^{-1}$ during the preincubation (30 min) and then removed by washing. To achieve satisfactory signal-to-noise ratio, an individual diffusion measurement of diI often involved adding several successive recoveries from repeated bleaches at a single site. The immobile sites are those which showed less than 10% recovery after photobleaching. The diffusion coefficient (D) and mean fractional recovery are shown as mean $\pm$ s.d. and were calculated using the theory developed by Axelrod *et al.*⁵. D is related to the half-time for recovery, $t_{1/2}$, by $D = \gamma\omega^2/(4t_{1/2})$ where γ is a factor that takes into account the extent of bleaching and ω is the focused laser beam radius at the e^{-2} intensity point. The fractional recovery is the per cent of fluorescence that returns after bleaching of a site. The number of sites measured under each treatment is given in parentheses.

incubated with 50 mM α -methyl mannoside, demonstrating that the Con A was binding to sites on the surface of the cells and had not been internalized.

Most Con A receptors were mobile on control cells (Fig. 1), with a diffusion coefficient of $1.3 \times 10^{-10} \text{ cm}^2 \text{ s}^{-1}$, comparable with that measured for membrane proteins in many cells⁴. In control cells neither the number of immobile sites nor their diffusion coefficient changed for periods of growth of up to 2 weeks (by which time cells had reached confluency).

However, during differentiation induced by DMSO, we found a dramatic reduction in the number of sites at which receptors were mobile (Fig. 1). This effect was seen within 36 h of the beginning of treatment and persisted throughout the 2-week study period, by which time the cells had extended long processes. There was no accompanying segregation of Con A receptors to either soma or processes, nor, as shown in Table 1, did mobile sites persist in any particular region of the cell. It is clearly possible that in conditions that leave most receptors mobile, more subtle changes in the diffusion coefficient might be found on different parts of the cell⁶. WGA receptors were also distributed over the entire cell surface but showed no evidence of a change in the number of immobile sites during differentiation (Table 1). Treatment of cells with 1 mM dibutyl cyclic AMP for 1 week also induces morphological and electrophysiological maturation², and caused a similar degree of immobilization of Con A binding sites. These effects are unlikely to be related

simply to the lowered rate of cell division seen with differentiating agents because de Laat *et al.* have shown protein diffusion to be slowest during mitosis⁷. Furthermore, serum withdrawal markedly slowed cell division⁸ and induced neurite extension but did not reduce the mean fractional recovery or increase the number of immobile sites (all sites were mobile, $n = 6$).

Further investigation suggested that the cytoskeleton contributed to the developmental immobilization. Although restrictions of mobility might conceivably derive from changes in 'fluidity' of membrane lipid^{9,10}, the diffusion coefficient of the lipid probe diI was not significantly different in control and DMSO-treated cells (Table 1), suggesting that changes in the lipid environment were not responsible. On the other hand, as shown in Table 1, dissociation of components of the cyto-

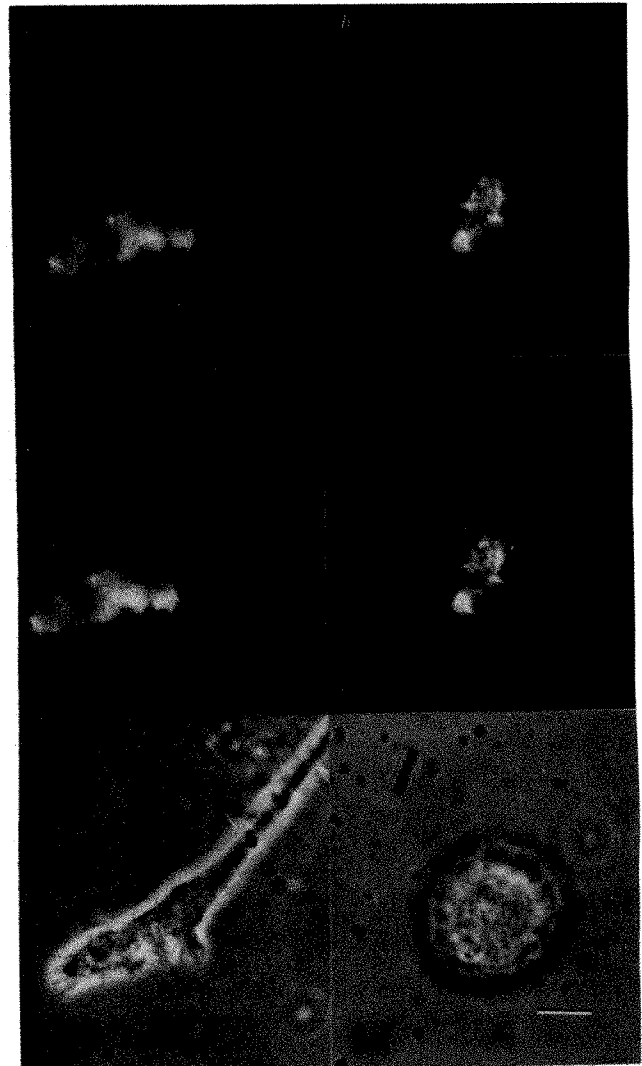


Fig. 2 Growth cone (a) and cell body (b) of a neuroblastoma cell treated with 2% DMSO, washed in balanced salt solution, labelled with succinyl Con A, fixed for 30 min in 3.7% formaldehyde and then labelled with $0.5 \mu\text{g}$ of NBD-phalloidin in 0.5 cm^3 of balanced salt solution for 30 min and washed twice. It shows the correlation of the internal actin network, labelled with the NBD-phalloidin (top) and the external Con A binding sites (middle). Bright field is shown at the bottom. In other experiments cells were labelled singly with either NBD-phalloidin (after fixation) or Con A (both before and after fixation) alone and the same patterns were observed. The NBD-phalloidin did not label unfixed cells, even if incubated for 48 h. Fluorescence photomicrographs were taken using epifluorescent illumination and standard Zeiss filters for rhodamine or fluorescein (for the NBD-phalloidin). Control experiments revealed no 'bleed through': rhodaminated succinyl Con A was not visible at the NBD excitation wavelength, and the NBD was not visible at the rhodamine excitation wavelength. Scale bar, $10 \mu\text{M}$.

skeleton reversed the developmental immobilization. Most dramatic was the effect of cytochalasin B, an agent that disrupts microfilaments, and which reduced the percentage of immobile sites in differentiated cells from 79 to 23%. Incubations with colcemid for 30 min reversibly disrupt microtubules in these cells¹¹ but only slightly increased the mobility of Con A receptors, suggesting a lesser role for microtubules in the immobilization process. Prolonged treatment (90 min) did increase the mobile fraction, but cells in these conditions had lost all processes and many floated free. WGA receptors are immobilized by cytochalasin-insensitive mechanisms on both differentiated and undifferentiated cells. Thus, there are apparently constraints in the mobility of some surface molecules which are distinct from those which act on Con A receptors.

We pursued this suggestion of cytoskeleton-membrane interaction by correlating the distribution of internal actin with that of Con A receptors. A fluorescent derivative of the mushroom toxin phalloidin has recently been demonstrated to bind actin specifically and with high affinity¹². This 7-nitrobenz-2-oxa-1,3-diazole phalloidin (NBD-phalloidin) does not seem to alter the distribution of actin filaments (by comparison with staining with anti-actin antibodies)¹³ even though the closely related toxin phalloidin can cause conversion of G- to F-actin¹⁴. The actin network, as revealed by NBD-phalloidin, is concentrated at the cell periphery and growing tips (Fig. 2). This punctate peripheral distribution was similar in control and DMSO-treated cells, and is identical to the distribution of actin described using anti-actin antibodies¹⁵. Photographs of the ConA receptors on the same cell in the same focal plane show the remarkable correlation of Con A binding sites with regions of actin accumulation. This correlation was noted on surfaces on the cell soma, along the processes and on the growth cone. WGA binding sites, on the other hand, often were not associated with actin.

Interactions between actin and cell membrane receptors are thought to underlie capping in lymphocytes and other cells¹⁶⁻¹⁸ and to guide anisotropic movements on cell surfaces¹⁹. In some systems, there is evidence for linkage proteins that mediate cytoskeleton-membrane interactions²⁰. During morphological differentiation of neurones, the immobilization of selected cell-surface components, under the influence of genetic and environmental cues, may underlie acquisition of structural asymmetry and could serve not only to guide directed axon outgrowth, but also to enhance anatomical registration of pre- and postsynaptic structures.

We thank L. Barak for his gift of NBD-phalloidin, A. Waggoner for his gift of diI and M. Nirenberg, R. Blumenthal and P. G. Nelson for their support and the use of their facilities.

Received 20 October 1980; accepted 11 February 1981.

- Kimhi, Y., Palfrey, C., Spector, I., Barak, Y. & Littauer, U. Z. *Proc. natn. Acad. Sci. U.S.A.* **73**, 462-466 (1976).
- Chalazonitis, A., Minna, J. D. & Nirenberg, M. *Exp. Cell Res.* **105**, 269-280 (1977).
- Nelson, P., Christian, C. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 123-127 (1976).
- Elson, E. L. & Schlessinger, J. in *The Neurosciences, Fourth Study Program* (eds Schmitt, F. O. & Worden, F. G.) 691-702 (MIT Press, 1979).
- Axelrod, D., Koppel, D., Schlessinger, J. & Elson, E. *Biophys. J.* **16**, 1055-1069 (1976).
- de Laat, S. W., Van der Saag, P., Elson, E. & Schlessinger, J. *Biochem. biophys. Acta* **558**, 247-250 (1979).
- de Laat, S. W., Van der Saag, P., Elson, E. & Schlessinger, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1526-1528 (1980).
- Seeds, N. W., Gilman, A., Amano, T. & Nirenberg, M. *Proc. natn. Acad. Sci. U.S.A.* **66**, 160-167 (1970).
- Axelrod, D., Wight, A., Webb, W. & Horwitz, A. *Biochemistry* **17**, 3604-3609 (1978).
- Frye, L. D. & Edidin, M. *J. Cell Science* **7**, 319-335 (1970).
- Spiegelman, B. M., Lopata, M. A. & Kirschner, M. W. *Cell* **16**, 253-263 (1979).
- Barak, L. S., Yocum, R., Nothnagel, E. & Webb, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 980-984 (1980).
- Verderame, M., Alcorta, D., Egnor, M., Smith, K. & Pollack, R. *Proc. natn. Acad. Sci. U.S.A.* **66**, 6624-6628 (1980).
- Wieland, T. *Naturwissenschaften* **64**, 303-309 (1970).
- Marchisio, P. C., Osborn, M. & Weber, K. *J. Neurocytol.* **7**, 571-582 (1978).
- de Petris, S. in *Dynamic Aspects of Cell Surface Organization* Vol. 3 (eds Poste, G. & Nicolson, G. L.) 643-728 (North-Holland, Amsterdam, 1977).
- Karnovsky, M. J. & Unanue, E. R. *Fedn Proc.* **32**, 55-59 (1973).
- Schindler, M., Koppel, D. E. & Scheetz, M. P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1457-1461 (1980).
- Smith, B. A., Clark, W. R. & McConnell, H. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5641-5644 (1979).
- Geiger, B. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4127-4131 (1980).

Role of cyclic AMP in a serotonin-evoked slow inward current in snail neurones

Philippe Deterre*, Danièle Paupardin-Tritsch*, Joël Bockaert† & Hersch M. Gerschenfeld*

* Laboratoire de Neurobiologie, Ecole Normale Supérieure, 46 Rue d'Ulm, 75230 Paris Cedex 05, France

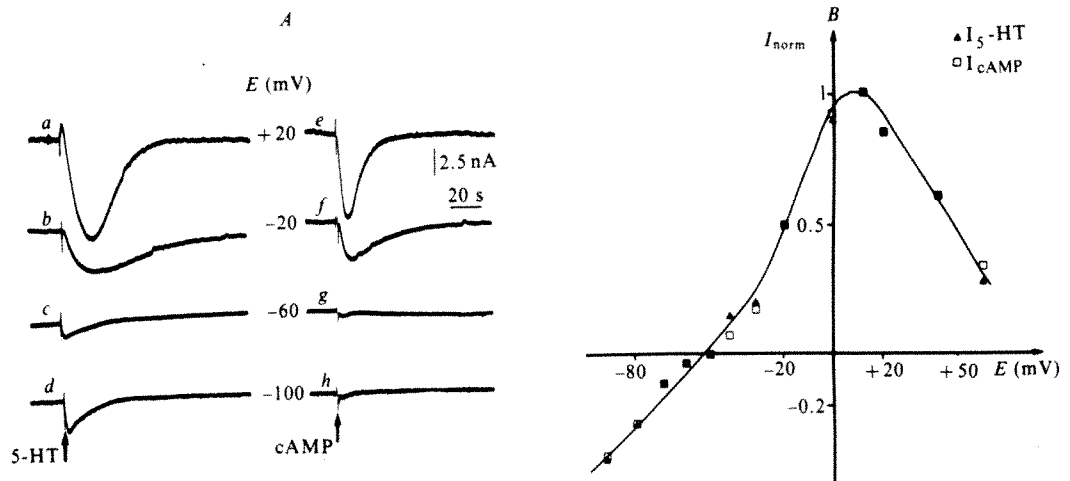
† Laboratoire de Physiologie Cellulaire, Collège de France, 11 Place Marcellin Berthelot, 75230 Paris Cedex 05, France

One model of synaptic transmission suggests¹ that transmitters modify postsynaptic permeability through the intermediary of cyclic AMP. Thus, serotonin (5-hydroxytryptamine) evokes in molluscan neurones a decrease in a voltage-dependent K^+ conductance^{2,3} which in turn generates a slow inward current when studied in steady voltage-clamp conditions³. The serotonin-induced increase of the plateau phase of the spike of an *Aplysia* sensory neurone can be mimicked by both intracellularly injected cyclic AMP and extracellularly applied phosphodiesterase inhibitors⁴, suggesting that cyclic AMP mediates the effect. We have tested whether a similar mechanism could account for the serotonin slow inward current in identified snail neurones and have found that the intracellular injection of cyclic AMP, but not of cyclic GMP or 5'-AMP, evokes a slow inward current showing similar voltage dependence, inversion potential and ionic properties to the serotonin slow inward current. Phosphodiesterase inhibitors at low concentrations (1-20 μ M) potentiate the serotonin slow inward current and at higher concentrations evoke by themselves an inward current, partially or totally occluding the serotonin and cyclic AMP currents. Finally, we have found that in homogenates of pooled identified snail neurones serotonin stimulates the adenylate cyclase, increasing its activity by 50-100%.

The experiments were carried out on identified neurones of the viscerio-abdominal ganglionic mass of *Helix aspersa* (cells F5, F6, E1, E2, E15, E17 of Kerkut *et al.*⁵). The ganglia were isolated in a chamber and continuously bathed in circulating saline of the following composition in mM: NaCl 120, KCl 5, CaCl₂ 6, MgCl₂ 3.5, HEPES 10 at pH 7.4. The identified neurones were impaled with double-barrelled microelectrodes connected to a voltage-clamp circuit⁶. Serotonin was applied iontophoretically or through the circulating bath solution. Tetraethylammonium (TEA) ions, Cs⁺, EGTA, cyclic AMP, cyclic GMP and 5'-AMP were injected intracellularly by passing appropriate inter-barrel currents⁷ between the barrels of double microelectrodes. One of the barrels was filled with TEA-Cl (1 M), CsCl (1 M), EGTA (0.5 M) or a 0.1 M solution of a monosodium compound of one of the nucleotides and the other with 3 M KCl. In the experiments in which both cyclic AMP and a channel blocker (or EGTA) were intracellularly injected by passing inter-barrel currents in the same cell, assemblies of three intracellular pipettes were used. NaCl was replaced in the saline by either Tris-Cl or sucrose, keeping the molarity constant. CaCl₂ in the saline was replaced by CoCl₂. When the extracellular Ca²⁺ concentration [Ca²⁺]_o was increased or TEA-Cl added to the saline, the molarity was kept constant by reducing the NaCl in the medium. Phosphodiesterase inhibitors (isobutyl methylxanthine, IBMX; RO 20-1724) were added to the circulating saline. Adenylate cyclase activity was assayed in pools of 40 identified snail neurones dissected out from 15 ganglia, according to the technique of Levitan⁸ slightly modified.

Figure 1A compares the effects of extracellular application of serotonin and the intracellular injection of cyclic AMP in the same snail neurone. When the membrane potential was held between -100 and -30 mV, serotonin evoked a biphasic inward

Fig. 1 A, Voltage-clamped snail neurone. The iontophoretic applications of serotonin (5-HT) (a-d) and the intracellular injections of cyclic AMP (cAMP) (e-h) were performed in sequence while holding the cell at different potential levels (centre). Note that the early small outward current evoked by serotonin in a (due to the activation of other receptors⁹) is not present in the cyclic AMP response (d). **B,** *I-V* curve relating the net currents evoked by iontophoretic applications of serotonin (▲) and intracellular injections of cyclic AMP (□) in a snail neurone bathed in a saline containing 30 mM TEA to suppress other components of the membrane conductance and of the serotonin response. The steady-state *I-V* curves were first obtained by stepping the voltage from a holding level (+10 mV) to different levels of potential between -90 and +50 mV and measuring the current 1 s after the onset of the pulse. The net currents evoked by serotonin and cyclic AMP were obtained by subtracting the steady-state curves obtained with serotonin and cyclic AMP from the control steady-state curves.



current (Fig. 1A, c, d) which, as previously shown^{3,9}, was associated with an increase in Na^+ conductance. At levels more depolarized than -30 mV (Fig. 1A, a, b), serotonin evoked a much slower inward current which showed a latency of 3-5 s. In contrast, no effect of intracellularly injected cyclic AMP was observed when the cell potential was held between -30 and -100 mV (Fig. 1A, g, h), but at potentials more depolarized than -30 mV (Fig. 1A, e, f) cyclic AMP also evoked a slow inward current whose amplitude increased with further

depolarization (Fig. 1A, e) in parallel with the serotonin response. The only difference between the two responses was that the cyclic AMP response never showed other components resulting from the activation of other serotonin receptors such as in Fig. 1A, a.

The current-voltage (*I-V*) curve of Fig. 1B relates the normalized values of the net currents evoked by either extracellularly applied serotonin (full triangles) or intracellularly injected cyclic AMP (squares) to the levels at which the voltage was stepped for measuring the currents. It is clear that the same *I-V* curve fits perfectly the values of the net currents evoked by either serotonin or cyclic AMP. Both currents show a maximal value at about +10 mV and an inversion potential at about -55 mV. Even if the inversion potential of both responses differed from the potassium potential, E_K (-75 mV in the snail neurones bathed in the same saline), it was sensitive to changes in the extracellular K^+ concentration. Intracellular injection of either 5'-AMP or cyclic GMP inside the same identified snail neurones had no effect.

The fact that a cyclic AMP-induced slow inward current was only seen in cells that exhibited a serotonin slow inward current, together with the similarity of the *I-V* curves of the two slow inward currents, suggests that these currents reflect the same conductance change. To test this hypothesis we compared the ionic sensitivity and pharmacology of the two currents. Neither the serotonin response nor the cyclic AMP response was affected by the replacement of two-thirds of the extracellular NaCl by either Tris-Cl or sucrose or by the presence of 30 mM TEA in the extracellular medium. The amplitude of both responses increased in parallel when $[\text{Ca}^{2+}]_o$ was increased up to 20 mM. Moreover, replacement of the $[\text{Ca}^{2+}]_o$ with Co^{2+} generated by itself an inward current and occluded the effects of both serotonin and cyclic AMP while intracellular injection of the K^+ -channel blockers TEA and Cs^+ , which evoked an inward current by themselves, also blocked the slow currents evoked by both serotonin and cyclic AMP. As was also the case for the serotonin slow inward current, EGTA intracellular injection had no blocking effect on the cyclic AMP slow inward current. Finally, when serotonin application evoked a maximal slow inward current it occluded any further effect of cyclic AMP and vice versa. Table 1 compares the properties of both slow inward currents.

These results indicate that the extracellular application of serotonin and the intracellular injection of cyclic AMP both induce slow inward currents which reflect a decrease in the same voltage-dependent and $[\text{Ca}^{2+}]_o$ -sensitive K^+ conductance.

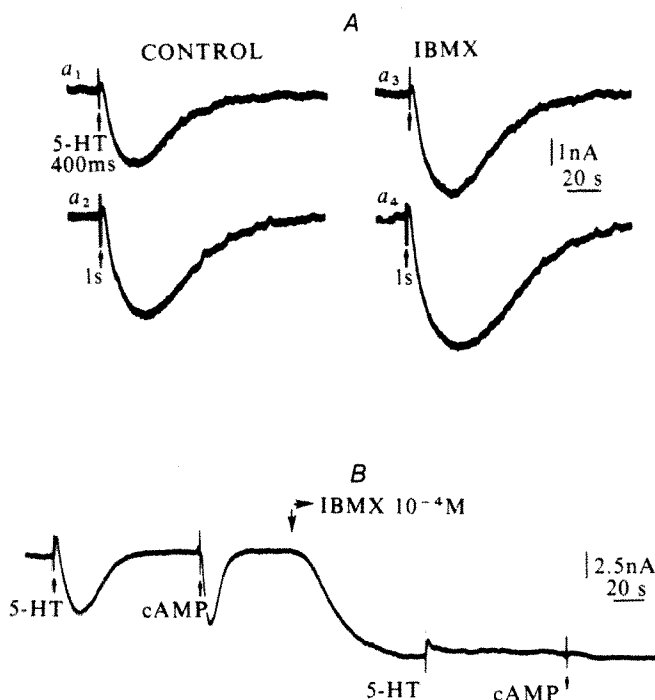


Fig. 2 A, Potentiating effect of IBMX on the serotonin slow inward current. a_1 , a_2 , Control serotonin slow inward currents obtained with the same iontophoretic current by varying the duration of the iontophoretic pulse a_3 , a_4 Responses to the same iontophoretic serotonin pulses as in a_1 , a_2 but in the presence of 20 μM IBMX in the extracellular medium. **B,** In another snail neurone, extracellular application of 100 μM IBMX evokes by itself an inward current and occludes the effects of the application of serotonin and cyclic AMP.

The next question is whether cyclic AMP mediates the effects of serotonin on the K^+ conductance whose decrease gives rise to the serotonin slow inward current. If a causal relationship exists between serotonin and cyclic AMP effects, it could be expected that the inhibitors of phosphodiesterase, the enzyme involved in the inactivation of cyclic AMP, should potentiate the response to serotonin. The results of our experiments were somewhat more complex. When IBMX was added to the bath in a concentration which, according to the cell, varied between 1 and 20 μ M, it caused, as expected, an increase in both amplitude and duration of the serotonin slow inward current (Fig. 2A, compare a_1 with a_3 and a_2 with a_4). However, at higher doses (up to 100 μ M), IBMX itself evoked an inward current which occluded both the serotonin and cyclic AMP responses (Fig. 2B). Similar effects were observed when using other methylxanthine-type phosphodiesterase inhibitors such as theophylline and with non-xanthine compounds such as RO 20-1724.

Our experiments suggest that the serotonin slow inward current results from the stimulation of an adenylate cyclase by

Table 1 Comparative properties of slow inward currents evoked by serotonin I_{S-HT} and cyclic AMP I_{cAMP}

	I_{S-HT}	I_{cAMP}
Increase of $[Ca^{2+}]_0$	Increase	Increase
Decrease of $[Ca^{2+}]_0$	Decrease	Decrease
Co^{2+}	Block	Block
Intracellular TEA	Block	Block
Intracellular Cs	Block	Block
Intracellular EGTA	No effect	No effect

serotonin. The latency of the serotonin response and the potentiating effects of the phosphodiesterase inhibitors favour such a view because it can be assumed that there is a basal activity of the cyclase and that the higher concentrations of IBMX could evoke increased accumulation of cyclic AMP which in turn would evoke a maximal decrease in K^+ conductance and thus occlude the responses to further extracellular application of serotonin or intracellular injection of cyclic AMP.

This hypothesis was supported by assays of adenylate cyclase activity. The adenylate cyclase activity of homogenates of pooled identified snail neurones, measured by the production of cyclic AMP from ATP during 30-min periods, was 151 ± 12 fmol (mean $\pm$ s.e.m. for triplicates) in basal conditions and 283 ± 4 fmol in the presence of 10 μ M serotonin, that is, serotonin increased the enzyme activity by 87%.

In conclusion, the present results suggest that the decrease in voltage-dependent K^+ conductance associated with the slow inward current evoked by serotonin in snail neurones is actually due to serotonin stimulation of the adenylate cyclase activity. In a larger context these results, together with other effects of serotonin on molluscan neurones probably mediated by cyclic AMP^{6,10}, support the idea that some synaptic actions may also depend on the activation of a second messenger system¹.

This work was supported by grants from the CNRS, DGRST and INSERM, France. P.D. was supported by a DGRST fellowship.

Received 31 October 1980; accepted 18 February 1981.

- Greengard, P. *Nature* **260**, 101–108 (1976).
- Klein, M. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6912–6916 (1980).
- Paupardin-Tritsch, D., Deterre, P. & Gerschenfeld, H. M. *Brain Res.* (in the press).
- Klein, M. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3512–3516 (1978).
- Kerkut, G. A., Lambert, J. C., Gayton, R., Loker, J. E. & Walker, R. J. *Comp. Biochem. Physiol. A* **50**, 1–25 (1975).
- Ginsborg, B. L. & Kado, R. T. J. *Physiol., Lond.* **245**, 713–725 (1975).
- Eccles, J. C., Eccles, R. M. & Ito, M. *Proc. R. Soc. B* **160**, 197–211 (1964).
- Levitan, I. B. *Brain Res.* **154**, 404–408 (1978).
- Gerschenfeld, H. M. & Paupardin-Tritsch, D. J. *Physiol., Lond.* **243**, 427–453 (1974).
- Drummond, A., Benson, J. A. & Levitan, I. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5013–5017 (1980).

Sertoli–Leydig cell communication via an LHRH-like factor

R. M. Sharpe*, H. M. Fraser*, I. Cooper* & F. F. G. Rommerts†

*MRC Reproductive Biology Unit, Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh EH3 9EW, UK

†Department of Biochemistry, Division of Chemical Endocrinology, Erasmus University, Rotterdam, The Netherlands

The primary function of the testosterone secreted by Leydig cells is the maintenance of spermatogenesis and hence fertility¹. This action of testosterone is mediated by the Sertoli cells which nourish and support the developing spermatozoa². As normal Sertoli cell function is so critically dependent on normal Leydig cell function, a regulatory influence of the Sertoli cells on the Leydig cells has been suggested^{3,4}. Indeed, follicle-stimulating hormone (FSH), which acts only on Sertoli cells⁵, can also cause profound changes in Leydig cell function^{6–8}, although how this is effected is unknown. We recently hypothesized that the Sertoli cell might be the source of a luteinizing hormone releasing hormone (LHRH)-like factor which we detected in the interstitial fluid surrounding the Leydig cells⁹. As injected LHRH agonists cause impairment of gonadal function^{10–13} and directly inhibit FSH-induced changes in Leydig cell function¹⁴ through specific membrane receptors^{15–17}, this 'LHRH-like' factor has all the correct credentials for the postulated messenger between the Sertoli and Leydig cells³. Here, we strengthen this case by demonstrating that seminiferous tubules from both the rat and the stump-tailed macaque (*Macaca arctoides*) contain a factor which has LHRH-like receptor-binding and biological activity *in vitro*, but which is immunologically distinct from native LHRH. We have also shown that this factor is secreted *in vitro* by cultured rat Sertoli cells.

Because of their ease of preparation and the specificity and affinity with which they bind radiolabelled LHRH agonists^{15–17}, isolated rat Leydig cells provide a simple, sensitive and precise radioreceptor assay for LHRH-like factors⁹. We have used this method to demonstrate that crude acid extracts of rat seminiferous tubules contain a factor which cross-reacts in parallel with unlabelled LHRH agonist (Fig. 1), whereas similar extracts of liver are inactive (data not shown). However, such crude extracts contained large amounts of protein (which may interfere in the receptor assay) and their potency decreased rapidly and progressively during storage at -20°C . An extract was therefore prepared which excluded proteins, based on methods for the extraction of native LHRH from the hypothalamus¹⁸. This acid/ethanol extract also gave parallel displacement curves to LHRH agonist in the receptor assay (Figs 1, 2), although more potent preparations tended towards non-parallelism when suppressing binding by 75% or more (see Fig. 3 and Fig. 1 legend). Acid/ethanol extracts of rat seminiferous tubules or stump-tailed macaque testis gave parallel displacement in the receptor assay and both stimulated ($P < 0.01$ and $P < 0.02$) the release of LH from rat hemipituitaries *in vitro* (Fig. 2).

Although the Sertoli cell seemed the most likely source of this LHRH-like factor, this was not certain as the seminiferous tubule preparation contained both germ cells and Sertoli cells, and was contaminated with some adherent Leydig cells. To resolve this problem we tested medium from Sertoli cell cultures. When assayed without dilution, this medium showed negligible LHRH-like activity in the receptor assay but after acid/ethanol extraction, which effectively concentrated it by a factor of 30, LHRH-like activity was clearly detected, two separate extracts giving dose-response curves parallel to LHRH agonist (Fig. 3). The specificity of this displacement can be judged by the fact that the Sertoli cell culture medium (SCCM)

extract and a rat seminiferous tubule extract gave only minor, although significant ($P < 0.05$), displacement in a Leydig cell human chorionic gonadotropin (HCG)-receptor assay which was run in parallel (Fig. 3). Interestingly, previous workers testing SCCM for inhibin-like activity reported the presence of a small molecule which stimulated pituitary LH release *in vitro*^{19,20}, and this probably corresponds to the LHRH-like factor which we have detected by receptor assay.

The nature of this LHRH-like factor produced by the Sertoli cells is unknown. It is clearly different from native LHRH, as rat seminiferous tubule extracts were largely inactive in an LHRH radioimmunoassay (Fig. 4). However, in a radioimmunoassay which we have developed using an antiserum raised against the LHRH agonist, seminiferous tubule extracts gave displacement

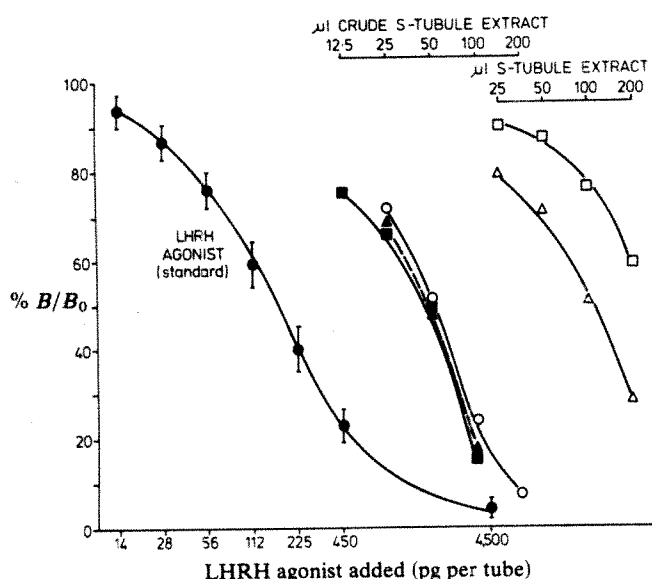


Fig. 1 Displacement of the binding of ^{125}I -LHRH agonist ((D-Ser-r-bu⁶, des-Gly-NH₂¹⁰) LHRH ethylamide) to dispersed Leydig cells by unlabelled LHRH agonist and rat seminiferous tubule (S-tubule) extracts. The methods used for the preparation of Leydig cells¹⁶ and the radioreceptor assay procedure were as described previously^{9,16}, except that the volumes used were 0.15 ml dispersed cells, 0.2 ml sample/standard and 0.05 ml ^{125}I -LHRH agonist; all samples and standards were run in duplicate. To prepare crude seminiferous tubule extracts, decapsulated testes from adult Sprague-Dawley rats were dispersed with collagenase²⁹, the isolated Leydig cells removed¹⁶ and the tubules washed three times with 50 ml 0.9% saline. After the last wash the tubules were centrifuged for 10 min at 1,000g, the supernatant discarded and the precipitated tubules homogenized in 0.1 M acetic acid (0.3 ml per g tissue) using a ground glass homogenizer. The homogenate was then centrifuged for 20 min at 1,000g, the supernatant aspirated and its pH adjusted to 7.5 using 1 M NaOH. This material was then assayed at multiple doses against unlabelled LHRH agonist; dilutions of standard and test material were made using 0.1 M acetic acid adjusted to pH 7.5 and containing 0.2% bovine serum albumin (fraction V, Sigma). In subsequent studies, an acid/ethanol extraction procedure was used, based on conventional techniques for the extraction of LHRH from the hypothalamus¹⁸. Thus, washed seminiferous tubules were homogenized in 0.1 M acetic acid (1 ml per g tissue) and the supernatant collected by centrifugation. The pellet was re-homogenized in the same volume of 0.1 M acetic acid and the resultant supernatant combined with the initial extract and lyophilized. The dried powder was then extracted with 0.1 M acetic acid in ethanol (1:40, v/v) at -20°C , using 0.5 ml per g original tissue. The extract was centrifuged for 10 min at 1,000g, the supernatant aspirated and evaporated to dryness in a stream of nitrogen at 45°C . The residue was dissolved in either 0.1 M acetic acid (pH 7.5) or 0.01 M phosphate-buffered saline (pH 7.5) at a concentration of 1 ml per 10 g original tissue. The pH of the extract was then adjusted to 7.5 using 1 M NaOH and the material stored at -20°C before assay. Representative displacement curves for three crude (○, △, ■) and two acid/ethanol (△, □) extracts are shown in relation to the LHRH agonist standard curve. The latter is a composite curve (mean $\pm$ s.d.) derived from 15 successive assays and illustrates its excellent reproducibility. Parallelism of standard and sample dose-response curves was tested using two-factor analysis of variance (with replication). Of a total of 9 crude and 12 acid/ethanol extracts assayed at 3 or more dilutions, only 5 samples gave dose-response curves which showed significant ($P < 0.05$ to $P < 0.01$) non-parallelism with the LHRH agonist standard; in all 5 instances, this deviation was apparent only at high concentrations of the extract which gave more than 75% suppression of binding (for example, see Fig. 3).

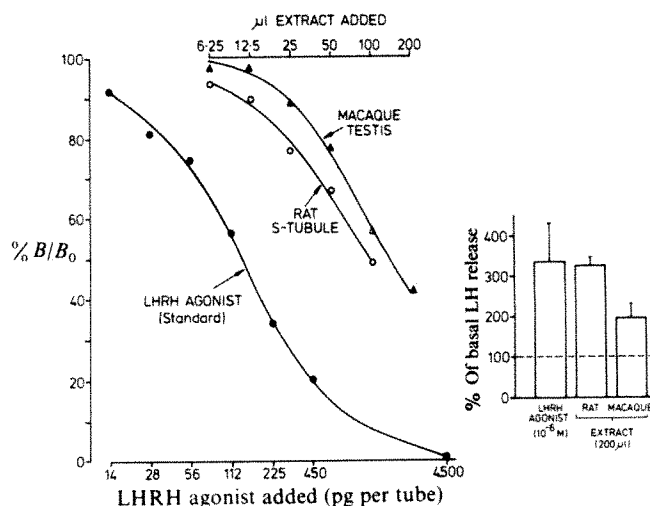


Fig. 2 *In vitro* receptor-binding activity and biological activity (inset) of acid/ethanol extracts of rat seminiferous tubules or macaque testis. Details of the extraction procedure for rat seminiferous tubules and the receptor assay method were as described in Fig. 1 legend. A single healthy, adult stump-tailed macaque of proven fertility was killed with an overdose of sodium pentobarbitone (Nembutal, May & Baker) and the testes immediately dissected out and kept on ice. After decapsulation the tissue was cut into ~1-g segments which were homogenized and extracted using acid/ethanol as described for rat seminiferous tubules; the final extract contained the equivalent of 10 g tissue per ml. The *in-vitro* biological activity of the extracts was determined by their ability to stimulate LH release from rat hemipituitaries⁹ using methods described elsewhere³⁰. One half of each donor hemipituitary was incubated in the presence of the extraction medium (pH 7.5) and the contralateral hemipituitary was incubated in the presence of 0.2 ml of either the extracts or LHRH agonist (10^{-6} M). LH release in the presence of the test materials has been expressed as a % of that released by the contralateral (control) hemipituitary. LH was determined by radioimmunoassay³¹ and results are the mean $\pm$ s.d. for three incubations.

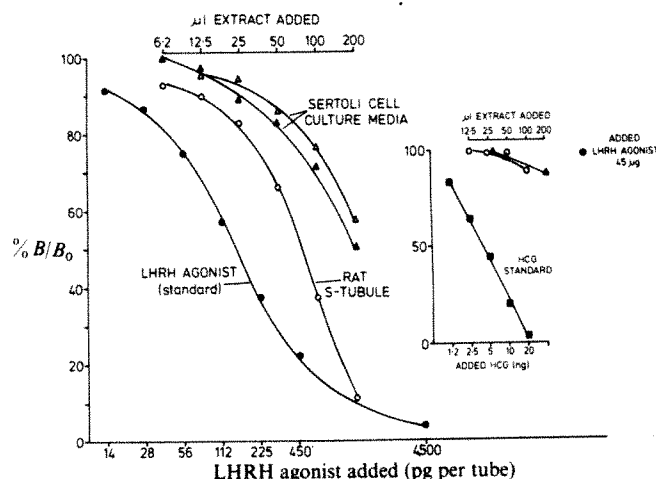


Fig. 3 Displacement of the binding of ^{125}I -LHRH agonist or ^{125}I -HCG (inset) to dispersed Leydig cells by acid/ethanol extracts of either rat seminiferous tubules or Sertoli cell culture medium. The same preparation and concentration of dispersed Leydig cells was used for both receptor assays. The HCG-receptor assay was performed as described elsewhere²⁹, and HCG CR115 (13,600 IU per mg) was used as standard. The seminiferous tubule extract was prepared as described in Fig. 1 legend, except that the pH of the initial acid extract was adjusted to 7.5, and the resulting precipitate removed by centrifugation, before lyophilization. Cultures of Sertoli cells from 21-day old rats were prepared as described elsewhere³² and maintained in culture for either 1 (▲) or 6 (△) days; in the latter instance the culture medium (5 ml per flask) was replaced after 1 and 5 days. Each culture flask contained ~700 μg protein and Sertoli and Leydig cells formed >90% and <3%, respectively, as judged by morphology, histochemistry^{32,33} and LH-dependent steroid production. Culture medium was stored at -20°C before extraction. The medium from at least eight separate cultures was pooled and lyophilized after adjusting the pH to 7.5. The powder was then extracted in acid/ethanol (1 ml per 5 ml culture medium) as described for seminiferous tubules in Fig. 1 legend. The final extract contained the equivalent of 30 ml culture medium per ml.

curves parallel to the unlabelled LHRH agonist, whereas native LHRH showed <0.1% cross-reaction (Fig. 4). As there is evidence that LHRH-like molecules may be produced in other extra-hypothalamic tissues²¹, for example the ovary^{22,23}, the availability of an antibody which does not measure hypothalamic LHRH may have widespread applications.

Impairment of the secretory function of Sertoli cells, as for example in cryptorchidism^{24,25}, is associated with marked hypertrophy of the Leydig cells²⁶, and similar morphological changes have been observed locally within the testis around implants of anti-androgens³. These data have been interpreted as evidence that the Leydig cells may be under the inhibitory control of an androgen-dependent factor from the Sertoli cells³. Various findings suggest that the LHRH-like factor from the Sertoli cell, which we have described, may be this inhibitor. First, all the direct effects of injected LHRH agonists on Leydig cell function are inhibitory²¹, and it is particularly relevant that

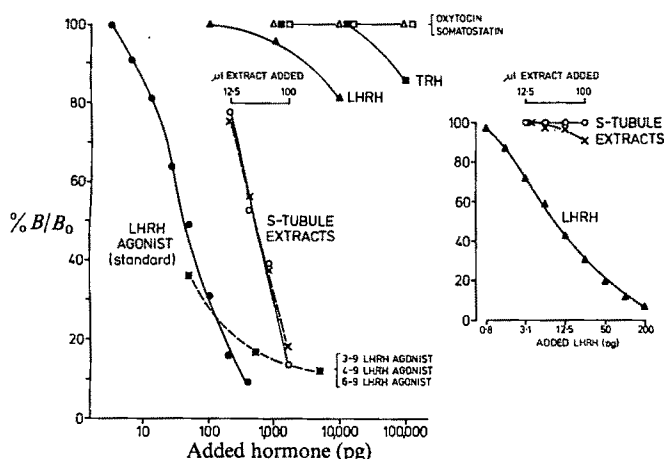


Fig. 4 Cross-reactivity of acid/ethanol seminiferous tubule extracts in radioimmunoassays for LHRH agonist and LHRH (inset). The radioimmunoassay method described previously for LHRH³⁴ was used for both assays and separation of bound and free hormone was achieved using ice-cold ethanol. ¹²⁵I-labelled LHRH or LHRH agonist was used as tracer in their respective assays and all samples and standards were run in duplicate. Details of the LHRH antibody and its specificity have been described elsewhere³⁵; LHRH agonist showed <0.1% cross-reaction in this LHRH radioimmunoassay. The antibody used in the LHRH agonist radioimmunoassay was obtained as follows. Five rabbits were immunized with ((D-Ser-t-bu⁶, des-Gly-NH₂¹⁰) LHRH ethylamide) (LHRH agonist) conjugated to human serum albumin using carbodiimide, as described for LHRH³⁶. Only one of the antisera (R103) cross-reacted with the seminiferous tubule extracts and this was used to develop a radioimmunoassay. The antibody was used at a final concentration of 1:12,000 and at this concentration specifically bound 26.2±4.6% (mean±s.d., n=4) of the added ¹²⁵I-LHRH agonist; the detection limit of the assay (the mass of standard required to suppress binding by 10%) was 7.5±1.1 pg. Of the peptides tested, only fragments of the LHRH agonist showed major cross-reaction with the antibody. TRH, thyrotropin releasing hormone.

the stimulatory effects of FSH on Leydig cell function (Sertoli cell-mediated effects) are totally abolished by concomitant treatment with LHRH agonists¹⁴. Second, treatment of rats with HCG, which markedly raises their intra-testicular levels of testosterone²⁷, stimulates the secretion of an LHRH-like factor into the interstitial fluid that surrounds the Leydig cells⁹, and this coincides temporally with the onset of Leydig cell down-regulation²⁸. The fact that changes as drastic as those detailed above can result from alteration of Leydig cell exposure to LHRH-like factors, suggests that this Sertoli cell hormone may have a central regulatory role in the maintenance of normal testicular function and fertility.

We thank Dr J. Sandow for helpful discussions and advice, Mrs M. Swaney for technical assistance, Drs J. Sandow and S. Dombey (Hoechst) for supplies of peptide hormones, and the National Pituitary Agency, NIAMDD, USA, for purified HCG and materials for rat LH radioimmunoassay.

Received 4 December 1980; accepted 11 February 1981.

- Hansson, V., Calandra, R., Purvis, K., Ritzén, M. & French, F. S. *Vitam. Horm.* **34**, 187-214 (1976).
- Means, A. R., Fakunding, J. L., Huckins, C., Tindall, D. J. & Vitale, R. *Recent Prog. Horm. Res.* **32**, 477-527 (1976).
- Aoki, A. & Fawcett, D. W. *Biol. Reprod.* **19**, 144-158 (1978).
- Rich, K. A., Kerr, J. B. & de Kretser, D. M. *Molec. cell. Endocr.* **13**, 123-135 (1979).
- Means, A. R. in *The Testis* Vol. 4 (eds Johnson, A. D. & Gomes, W. R.) 163-188 (Academic, New York, 1977).
- Odell, W. D. & Swerdloff, R. S. *Recent Prog. Horm. Res.* **32**, 245-288 (1976).
- Chen, Y. D. L., Payne, A. H. & Kelch, R. P. *Proc. Soc. exp. Biol. Med.* **153**, 473-475 (1976).
- Selin, L. K. & Moger, W. H. *Endocr. Res. Commun.* **4**, 171-182 (1977).
- Sharpe, R. M. & Fraser, H. M. *Nature* **287**, 642-643 (1980).
- Labrie, F. et al. *J. Androl.* **1**, 209-228 (1980).
- Rivier, C., Rivier, J. & Vale, W. *Endocrinology* **105**, 1191-1201 (1979).
- Sandow, J. & Hahn, M. *Acta endocr. copenh.* **88**, 601-610 (1978).
- Sharpe, R. M. & Fraser, H. M. *J. Reprod. Fert.* **60**, 359-368 (1980).
- Hsueh, A. J. W. & Erickson, G. F. *Nature* **281**, 66-67 (1979).
- Clayton, R. N., Katikineni, M., Chan, V., Dufau, M. L. & Catt, K. J. *Proc. natl. Acad. Sci. U.S.A.* **77**, 4459-4463 (1980).
- Sharpe, R. M. & Fraser, H. M. *Biochem. biophys. Res. Commun.* **95**, 256-262 (1980).
- Lefebvre, F. A., Reeves, J. J., Séguin, C., Massicotte, J. & Labrie, F. *Molec. cell. Endocr.* **20**, 127-134 (1981).
- Wheaton, J. E. & McCann, S. M. *Neuroendocrinology* **20**, 296-310 (1976).
- de Jong, F. H., Smith, S. D. & van der Molen, H. J. *J. Endocr.* **80**, 91-102 (1979).
- de Jong, F. H., Welschen, R., Hermans, W. P., Smith, S. D. & van der Molen, H. J. *J. Reprod. Fert. Suppl.* **26**, 47-59 (1979).
- Sharpe, R. M. *Nature* **286**, 12-14 (1980).
- Ying, S.-Y. & Guillemin, R. C. *hebd. Séanc. Acad. Sci. Paris D289*, 943-946 (1979).
- Ying, S.-Y. & Guillemin, R. C. *Abstr. Am. Endocrine Soc. Abstr.* **158** (1980).
- Hagenäs, L. & Ritzén, E. M. *Molec. cell. Endocr.* **4**, 25-34 (1976).
- Kerr, J. B., Rich, K. A. & de Kretser, D. M. *Biol. Reprod.* **21**, 823-838 (1979).
- Kerr, J. B., Rich, K. A. & de Kretser, D. M. *Biol. Reprod.* **20**, 409-422 (1979).
- Sharpe, R. M. *J. Reprod. Fert.* **55**, 365-371 (1979).
- Sharpe, R. M. *Biol. Reprod.* **22**, 851-858 (1980).
- Sharpe, R. M. & McNeilly, A. S. *Molec. cell. Endocr.* **18**, 75-86 (1980).
- Sharpe, R. M., Shahmanesh, M., Ellwood, M. G., Hartog, M. & Brown, P. S. *J. Endocr.* **65**, 265-273 (1975).
- Fraser, H. M. & Sandow, J. *J. Endocr.* **74**, 291-296 (1977).
- Rommerts, F. F. G., Kruger-Sewnarian, B. Ch., van Woerkom-Blik, A., Grootegeed, J. A. & van der Molen, H. J. *Molec. cell. Endocr.* **10**, 39-55 (1978).
- Rommerts, F. F. G., Kruger-Sewnarian, B. Ch., Grootegeed, J. A., de Jong, F. H. & van der Molen, H. J. *Acta. endocr. copenh.* **90**, 552-561 (1979).
- Jeffcoate, S. L., Fraser, H. M., Holland, D. T. & Gunn, A. *Acta. endocr. copenh.* **75**, 625-635 (1974).
- Jeffcoate, S. L., Holland, D. T. & Fraser, H. M. in *Physiological Effects of Immunity against Hormones* (eds Edwards, R. G. & Johnston, M. H.) 121-136 (Cambridge University Press, 1976).
- Fraser, H. M., Gunn, A., Jeffcoate, S. L. & Holland, D. T. *J. Endocr.* **61**, ix-x (1974).

Acceleration of β -receptor desensitization in combined administration of antidepressants and phenoxybenzamine

Fulton T. Crews

Department of Pharmacology and Therapeutics, University of Florida, Gainesville, Florida 32610, USA

Steven M. Paul & Frederick K. Goodwin

Clinical Psychobiology Branch, National Institute of Mental Health, Bethesda, Maryland 20205, USA

The delayed therapeutic effects of antidepressants (usually between 10 and 14 days)^{1,2} and of the tricyclic antidepressants in particular, are believed (on the basis of animal experiments³⁻⁸) to lie in a progressive decrease of the sensitivity of cortical β -adrenergic receptors. This is thought to be due to an increase in the synaptic concentration of noradrenaline, in turn accomplished by a decrease in the sensitivity of the presynaptic α_2 receptors which normally regulate noradrenaline secretion by a negative feedback mechanism. This model suggests that the desensitization of postsynaptic β -receptors by antidepressants should be accelerated by the inhibition of the presynaptic α_2 -adrenergic system, and we have indeed observed such an effect in preliminary studies with desipramine and phenoxybenzamine (PBZ) combined⁹. We now show that the administration of either tricyclic or monoamine oxidase inhibitor antidepressants in combination with PBZ, an irreversible α -adrenergic blocker^{9,10}, accelerates and intensifies the desensitization of β -adrenergic receptors. Our observations may have therapeutic implications.

Male Sprague-Dawley rats (125–150 g) were injected intraperitoneally (i.p.) with antidepressant, PBZ (7.5 mg per kg), combined antidepressant and PBZ (7.5 mg per kg), or vehicle, once daily. Rats were killed 24 h after the last dose of drug or vehicle. The cerebral cortices were dissected, weighed and homogenized using a Brinkmann polytron (setting 5, 50 s). Cortical membranes were sedimented by centrifugation (20,000g, 20 min at 4 °C), washed twice (20 vol w/v) and resuspended in 20 volumes of ice-cold 50 mM Tris-HCl buffer (pH 8.0). Aliquots of tissue equivalent to ~400–500 µg of protein were assayed for β -adrenergic receptors using ^3H -dihydroalprenolol (^3H -DHA) as described elsewhere¹¹. The binding of ^3H -DHA (0.1–3.0 nM, specific activity 32 Ci mmol⁻¹, NEN) was carried out at 23 °C for 20 min in a total incubation volume of 0.5 ml of 50 mM Tris-HCl buffer (pH 8.0). Assays were run in quintuplicate with two tubes containing 10 µM *dl*-propranolol to assess nonspecific binding. Incubations were terminated by rapid filtration at reduced pressure through Whatman GF/B glass fibre filters. Filters were washed five times with 5 ml of ice-cold buffer, dried, transferred to scintillation vials containing 10 ml of Aquasol (NEN) and counted in a liquid scintillation counter. Specific binding, defined as the total binding minus nonspecific binding, was ~85%.

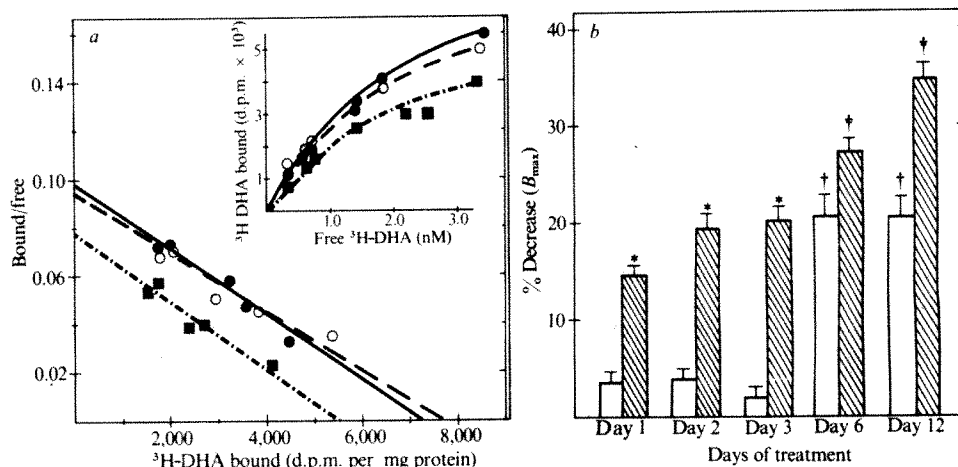
As in previous reports, acute treatment with desipramine (DMI, 7.5 mg per kg, i.p.) a prototype tricyclic antidepressant for 1, 2 or 3 days had no significant effect on the binding of ^3H -DHA to cerebral cortical membranes compared with vehicle-treated controls. Continuous administration of DMI for 6 or 12 days decreased ^3H -DHA binding by $24 \pm 4\%$ ($P < 0.05$) and $22 \pm 6\%$ ($P < 0.05$, Fig. 1), respectively. In contrast to DMI treatment alone, only one dose of combined DMI+PBZ significantly decreased ^3H -DHA binding ($14 \pm 2\%$, $P < 0.01$) compared with vehicle-treated controls (Fig. 1). Combined administration of DMI+PBZ for 2 and 3 days further reduced cerebral cortical β -adrenergic receptor density to that observed during chronic treatment with DMI alone (Fig. 1 and unpublished observations). Furthermore, during chronic administration of DMI+PBZ there was a significantly greater reduction in the density of cerebral cortical β -adrenergic receptors than that found during DMI administration alone (Fig. 1). For example, after 12 days of treatment with DMI and PBZ, specific ^3H -DHA binding was reduced by $34 \pm 3\%$ compared with $22 \pm 6\%$ with DMI alone ($P < 0.05$). All changes in ^3H -DHA binding after treatment with DMI or combined with DMI+PBZ, as determined by Scatchard analysis (Fig. 1), were due to a decrease in the density of β -adrenergic receptors and not a change in affinity (apparent K_d). These results indicate that combined DMI+PBZ administration desensitizes cortical β -receptors both more rapidly and to a greater extent than treatment with DMI alone.

Although there was a slight trend toward decreasing ^3H -DHA binding during chronic treatment with PBZ (7.5 mg per kg, i.p., data not shown) there were no statistically significant changes in ^3H -DHA binding after 1, 2, 3, 6, or 12 days of PBZ administration. The addition of PBZ (1 µM), DMI (1 µM) or combined DMI+PBZ (1 µM each) to cerebral cortical membranes *in vitro* did not significantly alter ^3H -DHA binding. As these concentrations are equivalent or even higher than those observed in washed membrane preparations, these results suggest that neither DMI nor PBZ directly interfere with ^3H -DHA binding.

To determine whether the synergistic actions of combined DMI+PBZ administration on cerebral cortical β -receptor desensitization were due to increased brain levels of desipramine, ^3H -DMI (7.5 mg per kg, i.p., 3.5 mCi mmol⁻¹, NEN) was injected into rats either alone or in combination with PBZ (7.5 mg per kg). Rats were killed 6 h later and the cerebral cortices dissected, weighed and homogenized in 20 volumes of alkaline ethanol. The radioactivity in an aliquot of the brain extract was determined in a liquid scintillation counter as described above. More than 90% of the radioactivity was extracted using this procedure. The levels of DMI and its metabolites were found to be 1.26 ± 0.09 ($n = 9$) and 1.14 ± 0.06 ($n = 9$) nmol per mg wet weight for rats treated with DMI alone or DMI in combination with PBZ, respectively. TLC of brain extracts from both groups of animals showed that 82% of radioactivity migrated with an R_F value identical to that of authentic DMI. These findings indicate that the effects of combined DMI+PBZ administration on cortical β -receptor desensitization are not due to changes in DMI metabolism.

To determine whether the rapid desensitization of cortical β -receptors found during combined PBZ+DMI (a prototype tricyclic antidepressant) treatment was true for other types of antidepressant drug, experiments were done using tranlylcypromine and pargyline, two chemically different monoamine oxidase inhibitors. Treatment with tranlylcypromine (10 mg per kg) alone for 1 or 2 days did not change ^3H -DHA binding. In contrast, combined tranlylcypromine (10 mg per kg)+PBZ (7.5 mg per kg) treatment reduced ^3H -DHA binding by $19 \pm 2\%$ ($P < 0.01$, Fig. 2) after 1 day of treatment and $19 \pm 1\%$ ($P < 0.01$) after 2 days of treatment. Six days of tranlylcypromine alone did decrease ^3H -DHA binding by $19 \pm 3\%$ ($P < 0.05$) compared with a decrease of $23 \pm 2\%$ ($P < 0.01$) with combined tranlylcypromine+PBZ treatment for 6 days. Pargyline (25 mg per kg), like tranlylcypromine, also had no significant affect on ^3H -DHA binding after 1 or 2 days of administration whereas after 6 days ^3H -DHA binding was reduced by $18 \pm 2\%$ ($P < 0.01$, Fig. 2). Combined pargyline (25 mg per kg)+PBZ (7.5 mg per kg) administration reduced ^3H -DHA binding by 5% after 1 day, 19% ($P < 0.01$) after 2 days and 23% ($P < 0.01$, Fig. 2)

Fig. 1 *a*, Scatchard analysis of ^3H -DHA binding to cerebral cortical membranes from rats treated with vehicle, DMI (7.5 mg per kg), or DMI (7.5 mg per kg)+PBZ (7.5 mg per kg) once daily for 2 days. The inset shows the saturation isotherm of ^3H -DHA binding as a function of increasing concentrations of ^3H -DHA. Bound/free, ratio of specifically bound to free ^3H -DHA. *b*, Effects of DMI (7.5 mg per kg) alone (open bars) or combined DMI (7.5 mg per kg)+PBZ (7.5 mg per kg) (hatched bars) on the development of cerebral cortical β -adrenergic receptor desensitization. Animals were administered vehicle or drugs i.p. once daily, and the number of β -adrenergic receptors determined 24 h after the last dose as described in the text. Values are the $\bar{x} \pm \text{s.e.m.}$ of the per cent change in the density (B_{max} as estimated by Scatchard analysis) of ^3H -DHA binding sites in cerebral cortical membranes from drug-treated animals compared with vehicle-treated controls. Control values were ~0.16 pmol ^3H -DHA bound per mg protein. Treatment with PBZ (7.5 mg per kg) alone had no significant effect on binding. Each experiment consisted of groups of six animals and was repeated twice. * $P < 0.01$; † $P < 0.05$; ‡ $P < 0.001$ (Student's *t*-test).



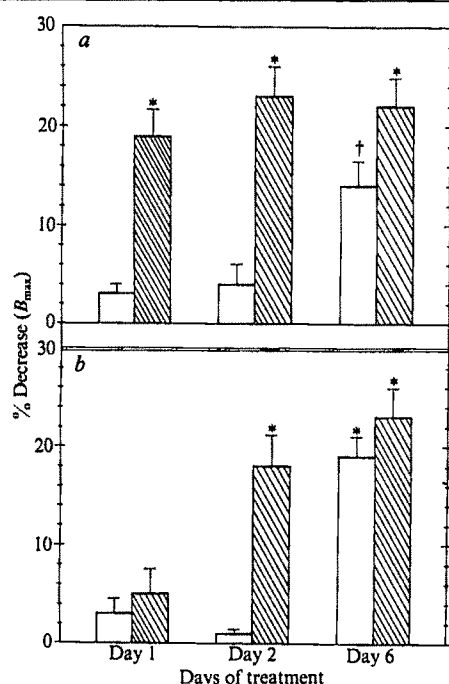


Fig. 2 *a*, Effects of tranylcypromine (10 mg per kg) alone (open bars) or combined tranylcypromine (10 mg per kg)+PBZ (7.5 mg per kg) (hatched bars) on the development of cerebral cortical β -receptor desensitization. *b*, Effects of pargyline (25 mg per kg) alone (open bars) or combined pargyline (25 mg per kg)+PBZ (7.5 mg per kg) (hatched bars) on the development of cerebral cortical β -receptor desensitization. Values are the $\bar{x} \pm s.e.m.$ of the per cent change in the density (B_{max} as estimated by Scatchard analysis) of ^3H -DHA binding sites in cerebral cortical membranes from drug-treated animals compared with vehicle-treated controls. Each experiment consisted of groups of six animals and was repeated twice. * $P < 0.01$; † $P < 0.05$ (Student's *t*-test).

after 6. In both these studies treatment with vehicle or PBZ (7.5 mg per kg) alone had no significant effect on ^3H -DHA binding. Thus, combined treatment of antidepressant monoamine oxidase inhibitors and PBZ can rapidly decrease β -adrenergic receptor density.

As the desensitization of cortical β -receptors during chronic tricyclic antidepressant treatment can be prevented by destruction of ascending noradrenergic neurones from locus coeruleus⁷, chronic treatment with tricyclic antidepressants probably increases the stimulation of β -receptors by increasing the synaptic concentration of noradrenaline. However, acute antidepressant administration decreases the adrenergic activity of locus coeruleus neurones apparently by increasing presynaptic α_2 -receptor inhibition¹². Recent studies in both the peripheral and central nervous systems have demonstrated that chronic antidepressant administration reduces this presynaptic α_2 -receptor inhibition^{8,13}, resulting in enhanced noradrenaline release which would be expected to desensitize postsynaptic cerebral cortical β -receptors. Thus, the increase in presynaptic α_2 -receptor inhibition during acute antidepressant treatment is opposite to the chronic and apparently therapeutically relevant effects of these agents.

In our experiments administration of PBZ probably blocks presynaptic α_2 -receptor feedback inhibition, allowing DMI, an inhibitor of noradrenaline uptake, to increase acutely the synaptic noradrenaline concentration and thereby rapidly desensitize cortical β -receptors. Consistent with this interpretation is the finding that yohimbine, another presynaptic α_2 blocker, in combination with DMI (5 mg per kg) decreased cortical β -receptor density in 4 days whereas DMI alone produced no change¹⁴.

The synergistic actions of tricyclic and monoamine oxidase inhibitor antidepressants with PBZ, an α -adrenergic antagonist, are probably due to increases in the synaptic concentrations of

noradrenaline. Rapid decreases in β -receptor binding sites have been shown to occur *in vitro* during exposure of cerebral cortical slices to β -adrenergic agonists¹⁵. The rapid decrease in ^3H -DHA binding reported here after DMI + PBZ treatment indicates that cerebral cortical β -adrenergic receptors can also be rapidly desensitized *in vivo*. The rapid desensitization of cerebral cortical β -receptors during administration of antidepressants and α -adrenergic antagonists has several important implications. As postsynaptic cortical β -receptors are thought to be directly involved in the therapeutic actions of most, if not all, antidepressants³, combined treatment with antidepressant and a presynaptic α_2 -receptor antagonist may provide a therapy with rapid onset of action. In addition, the greater reduction of β -receptor density produced by combined antidepressant + PBZ, compared with antidepressant alone, indicates that administration of an α -adrenergic antagonist may enhance the effects of antidepressants in previously nonresponsive patients. Finally, the decrease in β -receptor density after only one dose of combined antidepressant + PBZ treatment could serve as a rapid screening procedure for new antidepressant drugs.

We thank the Merrell-National Laboratories for desipramine and Smith Kline and French Laboratories for PBZ. F.T.C. thanks J. Axelrod for providing laboratory facilities and constructive comments.

Received 1 April 1980; accepted 6 February 1981.

- Oswald, I., Brezinova, V. & Dunleavy, D. L. *Br. J. Psychiat.* **120**, 673–677 (1972).
- Klein, D. F. & Davis, J. M. *Diagnosis and Drug Treatment of Psychiatric Disorders* (Williams & Wilkins, Baltimore, 1969).
- Sulser, F., Vetulani, J. & Mobley, P. L. *Biochem. Pharmac.* **27**, 257–261 (1978).
- Banerjee, S. P., King, L. S., Riggi, S. J. & Chanda, S. K. *Nature* **268**, 455–456 (1977).
- Sellinger, M. M., Frayer, A. & Mendels, J. *Fedn Proc.* **28**, 592 (1979).
- Vetulani, J., Stawary, R. J., Dingell, J. V. & Sulser, F. *Naunyn-Schmiedeberg's Archs Pharmacol.* **293**, 109–114 (1976).
- Wolfe, B. B., Harden, T. K., Sporn, J. R. & Molinoff, P. B. *J. Pharmac. exp. Ther.* **207**, 446–457 (1978).
- Crews, F. T. & Smith, C. B. *Science* **202**, 322–324 (1978).
- Paul, S. & Crews, F. T. *Eur. J. Pharmac.* **62**, 349–350 (1980).
- Cubeddu, L. X., Langer, S. Z. & Weiner, N. *J. Pharmac. exp. Ther.* **188**, 368–385 (1974).
- Bylund, D. B. & Snyder, S. H. *Molec. Pharmac.* **12**, 568–580 (1976).
- Nyback, H., Walters, J. R., Aghajanian, G. K. & Roth, R. H. *Eur. J. Pharmac.* **32**, 302–312 (1975).
- Svensson, T. H. & Usdin, T. *Science* **202**, 1089–1091 (1978).
- Johnson, R. W. *et al. Eur. J. Pharmac.* **67**, 123–127 (1980).
- Maggi, A., U'Prichard, D. C. & Enna, S. J. *Science* **207**, 645–647 (1980).

Isolation of a biologically active macrophage receptor for the third component of complement

Robert J. Schneider*†, Anthony Kulczycki Jr*‡, S. K. Law§ & John P. Atkinson*†

* Howard Hughes Medical Institute Laboratories and the Department of Medicine, † Division of Rheumatology and ‡ Division of Allergy and Immunology, Washington University School of Medicine, St Louis, Missouri 63110, USA
§ Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63110, USA

C3b, the major cleavage fragment of the third component of complement (C3), has been demonstrated to bind to a specific receptor on various mammalian cells¹. Tissue-bound C3b receptors have also been demonstrated, most conclusively in renal glomeruli². On interaction with C3b, C3b receptors are thought to initiate cellular functions such as phagocytosis and to determine the fate of complement-fixing soluble and particulate immune complexes. We report here the isolation by affinity chromatography of a macrophage glycoprotein with an apparent molecular weight (MW) of ~64,000 that has properties expected of the C3b receptor. It is a cell-surface macromolecule (labelled with ¹²⁵I and lactoperoxidase) which, in its isolated state, retains the ability to bind both C3 and C3b.

Human C3 was purified by the method of Tack and Prahl³, and was demonstrated to be haemolytically active³ and to retain its ¹⁴C-methylamine-binding capacity⁴. Three separate C3 preparations were used in the experiments; in most of them a preparation with >95% methylamine-binding capacity was used. C3b and C3d were prepared from C3 by trypsin digestion followed by gel filtration on Sephadex G-100 (ref. 5), and C3d was further purified by ion-exchange chromatography on DEAE-Sephacel⁴. The purities of the C3, C3b and C3d preparations were estimated by polyacrylamide gel electrophoresis (PAGE)⁶ to be at least 95%. C3 and C3b were coupled to Sepharose beads with cyanogen bromide with an efficiency of 57–86% and prepared for use as immunoadsorbents as described previously^{7–9}.

Elicited rabbit alveolar macrophages were collected, radiolabelled using modifications of the lactoperoxidase method¹⁰ and solubilized at 4 °C for 5 min in borate-buffered saline (BBS) containing 1% NP40 and 2 mM phenylmethylsulphonyl fluoride (PMSF)⁷. The 30,000g supernatant⁷ of the solubilized fraction of $\sim 2 \times 10^8$ cells was incubated with 0.4 ml of C3b-Sepharose for 1 h at room temperature. This mixture was then centrifuged at 300g for 5 min and the immunoadsorbent beads resuspended in 6.0 ml of BBS containing 1% NP40 and 2 mM PMSF, and washed twice in 0.7 × 4-cm glass columns with 4.0 ml of BBS containing 1%

NP40. C3b-binding molecules were subsequently eluted at room temperature with four successive 1.0-ml aliquots of 0.5 M acetic acid containing 1% NP40 into tubes with 0.375 ml of 2 M Tris, pH 8.6, containing 1% NP40 (to achieve complete, rapid neutralization). Using this procedure, we eluted 0.01–0.15% of the applied radioactivity from the immunoadsorbent columns.

When the acid eluate was analysed by SDS PAGE in rod gels in reducing conditions (1% β -mercaptoethanol), a single protein peak of MW $\sim 64,000 \pm 3,500$ (mean $\pm$ s.d. of 21 gels) was demonstrated (Fig. 1a). Purity, estimated by either determining the area under the curves or by cutting out and weighing the individual peaks, was routinely >90%. A similar peak was purified from the solubilized radiolabelled cells using C3b immunoadsorbent beads that had been prepared by the incubation of Sepharose with normal human serum¹¹, although other contaminating materials were also present (data not shown). No peak was observed if purification was attempted using beads prepared from heat-inactivated serum or using bovine serum albumin (BSA)-Sepharose. To exclude the possibility that trace contamination by IgG in our C3b preparation might allow co-purification of Fc γ receptor⁷, we preincubated the solubilized preparation with IgG-Sepharose. In three such experiments we recovered 92.5, 97.3 and 108% as much C3b receptor from the aliquot preincubated with IgG-Sepharose as from the aliquot not preincubated but isolated directly using only C3b-

Table 1 Rebinding of isolated C3b receptor to immunoadsorbents

Expt no.	Immunoadsorbent		Inhibitor protein	% Rebinding	% Specific rebinding	% Inhibition of specific rebinding
	For initial purification	For rebinding				
1	C3b	C3b	—	64.0 $\pm$ 1.1	60.6	—
		C3	—	60.8 $\pm$ 1.2	57.4	—
		IgG	—	7.1 $\pm$ 0.8	3.7	—
		BSA	—	3.4 $\pm$ 0.2	—	—
2	C3b	C3b	—	69.8 $\pm$ 2.0	62.3	None
		C3b	3.0 mg C3	21.8 $\pm$ 2.9	14.3	77.0
		IgG	—	7.5 $\pm$ 4.3	—	—
		C3b	—	18.7 $\pm$ 5.0	—	—
	IgG	IgG	—	57.8 $\pm$ 1.0*	—	—
		C3b	—	63.2 $\pm$ 2.0	56.1	—
	C3b†	IgG	—	7.1 $\pm$ 3.7	—	—
		C3b	—	18.0 $\pm$ 0.8	—	—
3	C3	IgG	—	51.9 $\pm$ 0.4	—	—
		C3b	—	31.6 $\pm$ 1.3	26.1	None
		C3b	3.0 mg OVAL	30.2 $\pm$ 2.0	24.8	5.0
		C3b	3.0 mg C3	8.1 $\pm$ 0.2	2.7	89.7
		C3b	3.0 mg IgG	29.1 $\pm$ 0.4	23.7	9.2
		IgG	—	5.4 $\pm$ 0.5	—	—
4	C3b	C3b	3.0 mg OVAL	27.1 $\pm$ 3.6	22.9	None
		C3b	0.3 mg C3b§	15.7 $\pm$ 1.7	11.5	49.8
		C3b	1.0 mg C3b	10.5 $\pm$ 1.9	6.3	72.4
		C3b	3.0 mg C3b	5.1 $\pm$ 0.7	0.9	96.1
		C3b	0.3 mg C3	19.3 $\pm$ 5.7	15.2	34.0
		C3b	1.0 mg C3	12.7 $\pm$ 2.5	8.5	63.0
		C3b	3.0 mg C3	6.3 $\pm$ 2.2	2.1	91.0
		IgG	3.0 mg OVAL	4.2 $\pm$ 4.5	—	—
5	C3b	C3b	3.0 mg OVAL	46.6 $\pm$ 1.6	41.5	None
		C3b	0.6 mg C3b§	20.2 $\pm$ 1.8	15.1	63.6
		C3b	0.6 mg C3d	49.3 $\pm$ 2.6	44.2	-6.6
		IgG	3.0 mg OVAL	5.1 $\pm$ 1.3	—	—

The ligand-binding activity of the putative receptor macromolecule was determined in a manner analogous to that described for the Fc γ receptor^{7,13}. Briefly, aliquots of the acid eluates were added to plastic tubes containing 0.10 ml immunoadsorbent. Inhibitor proteins were preincubated with the isolated receptor for 30 min at room temperature. Mixtures of immunoadsorbents and receptor preparations were incubated for 90 min at room temperature. Washed beads, supernatants and washes of triplicate samples were counted in a γ counter and the per cent rebinding of labelled receptor determined. Data presented are from representative experiments. The amounts of protein coupled to immunoadsorbents (mg protein per ml beads) were: C3b, 2.4–11.6 mg ml⁻¹; C3, 1.7–11.3 mg ml⁻¹; rabbit IgG, 6.8–15.7 mg ml⁻¹; and BSA, 4.5 mg ml⁻¹. For all rebinding studies, the amount of C3 or C3b present was in excess of that required to bind all of the available receptor.

* Represents rebinding of Fc γ receptor, which demonstrates the expected MW $\sim 35,000$ –55,000 (ref. 7) in SDS-polyacrylamide gels.

† The supernatant from the IgG-Sepharose column shown in the first part of expt 2 was adsorbed on C3b-Sepharose. The ligand-binding activity of the acid eluate from this second column is shown here.

‡ The supernatant from the C3b-Sepharose column shown in the first part of expt 2 was adsorbed on IgG-Sepharose. The ligand-binding activity of the acid eluate from this second column is shown here.

§ In these dose-dependent inhibition studies, ovalbumin (OVAL) was added to achieve a total protein content of 3.0 mg.

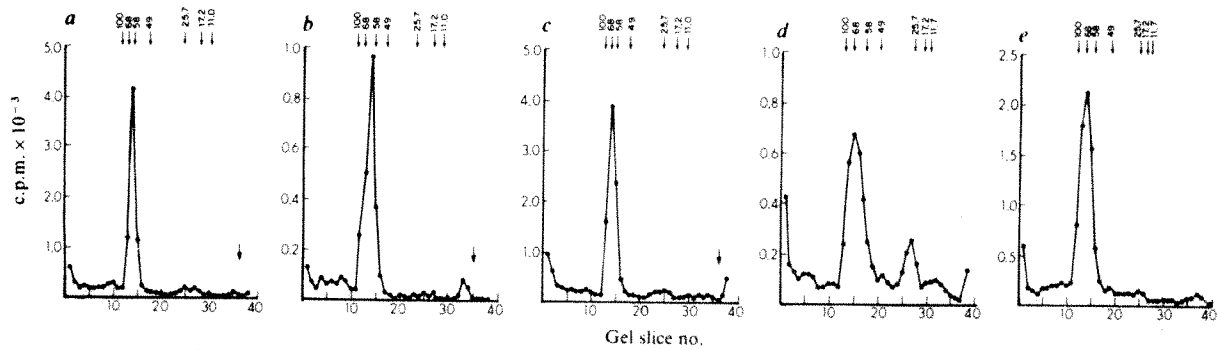


Fig. 1 SDS-PAGE in rod gels of C3b receptor. Representative results of five different types of experiment are shown. Samples were eluted with acid as described in the text and the eluates were prepared for electrophoresis using the Tris-glycine buffer system²¹ with 0.1% SDS and 9.0% acrylamide separating gels and 3.0% acrylamide stacking gels, using methods described previously⁷. Buffers for sample preparation contained 1% β -mercaptoethanol. The positions of six standard proteins are indicated by their molecular weights. Gels were sliced with an electric gel slicer, and radioactivity in 2-mm gel slices counted. *a*, Receptor from C3b-Sepharose. *b*, The ¹²⁵I-labelled, solubilized fraction was first incubated with IgG-Sepharose (to deplete Fc γ receptor from the preparation). This supernatant was subsequently incubated with C3b-Sepharose and the receptor was eluted and analysed as above. (The difference in counts between *a* and *b* reflects only the fact that one-fifth as many counts were loaded in the latter gel. As noted in the text, preincubation of the solubilized preparation with an IgG immunoadsorbent did not alter the quantity of C3b receptor obtained.) *c*, Receptor from C3-Sepharose. *d*, After collection by usual methods, the macrophages were resuspended in Hank's balanced salt solution (HBSS) at a concentration of 1.3×10^6 cells ml⁻¹. Cell suspension (8 ml) was placed on each of 45 100-mm plastic Petri dishes and incubated for 30 min at 37°C in a 5% CO₂ incubator. The nonadherent cells were removed by washing once with 5 ml of HBSS. The plated cells were solubilized with 4 ml of BBS containing 1% NP40 and 2 mM PMSF, and freed from the plates with a rubber policeman. The plates were serially washed with an additional 4 ml of the BBS containing 1% NP40 and 2 mM PMSF⁷. The solubilized preparation thus obtained was centrifuged at 30,000g for 20 min at 4°C and the C3b receptor isolated by affinity chromatography as above. A small peak at ~25,000 MW is present in this preparation (and also in the simultaneously prepared material obtained from the unplated cell population—not shown). Similar macromolecules were observed in 2 of the other 15 receptor preparations. *e*, Rabbit alveolar macrophages were collected and washed as above using sterile buffers and aseptic technique. Cells were subsequently suspended at 10^6 cells ml⁻¹ in sterile spinner flasks in minimal essential medium containing 10% fetal calf serum, 2 mM L-glutamine, 0.01% D-glucose, 1 mM pyruvate and 1:100 antibiotic/antimycotic mixture. Deoxyribonuclease I (10 μ g ml⁻¹) and ¹⁴C-glucosamine (1 mCi) were added to the medium. After 18–20 h culture at 37°C, the cells were collected by centrifugation and washed twice in 0.01 M phosphate, 0.15 M NaCl, pH 7.4 (phosphate-buffered saline, PBS) containing 1% BSA followed by a final wash in PBS. Cells were solubilized at 10^8 cells ml⁻¹ in PBS containing 1% NP40 and 2 mM PMSF⁷. The suspension was then centrifuged at 30,000g for 20 min at 4°C, and the supernatant applied to a C3b-Sepharose affinity column for 60 min at room temperature. The column was washed and eluted as described in the text to obtain receptor from intrinsically labelled cells.

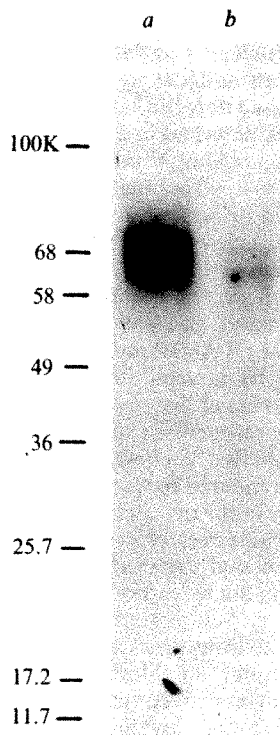


Fig. 2 SDS-PAGE in slab gels was performed according to the method of Laemmli⁶ using 0.1% SDS gels. Preparation of a linear gradient of polyacrylamide from 6 to 18% and a 3% stacking gel, gel fixation, staining, destaining and autoradiography have been described elsewhere¹³. Samples were prepared as described in Fig. 1 legend. *a*, Acid eluate from C3-Sepharose. *b*, A portion of the acid eluate in *a* was adsorbed on a new C3-Sepharose column. The rebound material was eluted with 6 M guanidine containing 1% NP40 and the sample prepared for analysis as above.

Sepharose. Macromolecules of identical MW were isolated from C3b-Sepharose and from C3-Sepharose (Fig. 1c); a similar protein was also obtained from an immunoadsorbent prepared by the coupling of C3 to Affi-Gel 10 (BioRad). The polyacrylamide gel electrophoretic mobility and pattern of the putative C3b receptor were not altered by: (1) addition of protease inhibitors (diisopropylfluorophosphate 10 mM, iodoacetamide 20 mM, or soybean trypsin inhibitor 2 mg ml⁻¹) in the solubilization buffer; (2) incubation temperature (4, 22 or 37°C) for affinity chromatography step; (3) omission of β -mercaptoethanol from the sample buffer; or (4) presence of Ca²⁺ and Mg²⁺ during solubilization and affinity chromatography. As alveolar cells contained an average of ~80% macrophages, adherent cells containing >95% macrophages⁷ were used in three experiments to obtain C3b receptor, and an identical macromolecule was isolated (Fig. 1d). Biosynthetic labelling with ¹⁴C-glucosamine demonstrated that, like the Fc ϵ ^{9,12} and Fc γ ⁷ receptors, the C3b receptor is a glycoprotein (Fig. 1e).

The ligand-binding activity of the affinity chromatography-purified material (always lower than the purity estimated by SDS-PAGE) was determined by the ability of the acid eluates to rebinding to a new immunoadsorbent column. In 12 such rebinding experiments the specific activity (rebinding to C3b-Sepharose minus rebinding to IgG-Sepharose or BSA-Sepharose) ranged from 22.9 to 62.3% (representative experiments are presented in Table 1). The per cent of rebinding was characteristic for each receptor preparation. The variability between preparations did not reflect differences in purity but rather partial denaturation during the obligatory acid elution procedure. The ligand-binding activity of the isolated receptor was independent of Ca²⁺ and Mg²⁺; it was stable for several hours at room temperature and at least several weeks at -70°C. The receptor rebound to both C3b-Sepharose and C3-Sepharose (Table 1, expt 1). Preincubation of the solubilized preparation with IgG-Sepharose, which removes Fc γ receptor⁷, did not alter the specific activity

of the C3b-binding protein eluted subsequently from a C3b-Sepharose column (Table 1, expt 2).

Two lines of evidence indicate that the ability of the isolated receptor to bind to both C3 and C3b was not due to conversion of C3 to either C3b or a C3b-like molecule during coupling to activated Sepharose. First, an alternative coupling procedure using Affi-Gel 10 allowed isolation of a similar macromolecule (see above). Second, both soluble C3 and soluble C3b were effective inhibitors of rebinding (Table 1, expt 3), although C3b was more effective in dose-dependent assays (Table 1, expt 4 is representative of three separate experiments). Soluble IgG, C3d and ovalbumin did not inhibit ligand-binding activity.

Aliquots of the acid eluates from the initial C3-Sepharose column were analysed in polyacrylamide gradient slab gels (Fig. 2a). Part of this material was rebound to a second C3-Sepharose column and then eluted with 6 M guanidine containing 1% NP40 (ref. 7), dialysed and analysed in gels (Fig. 2b). Electrophoretic patterns of both eluates demonstrated similar broad bands in the 58,000–68,000 MW region.

Like receptors for Fc γ ^{7,13} and Fc ϵ ^{9,12,14} previously isolated in these laboratories, the C3b-binding protein migrates as a broad band on polyacrylamide gels and retains specific ligand-binding activity. The molecular weight of the C3b receptor is higher than that of the Fc γ receptor isolated from rabbit alveolar macrophages, the latter having a MW of ~35,000–55,000 (ref. 7 and Table 1,*). Specific antibodies against analogously isolated Fc ϵ -binding molecules have been demonstrated to induce mast cell degranulation *in vitro*^{15,16}, providing definitive evidence that such binding molecules correspond to the biologically important Fc ϵ receptor. Similar evidence identifying Fc γ - or C3b-binding proteins conclusively as receptors has not yet been obtained. It now seems, however, that at least three important receptor molecules involved in immune responses are glycoproteins with apparent MWs in the 35,000–65,000 range and they can be isolated retaining ligand-binding activity in a highly purified state.

Although C3b was a more effective inhibitor in dose-dependent inhibition experiments than C3, both C3 and C3b bound to the putative complement receptor isolated from rabbit macrophages. Solubilization could possibly affect the observed specificity of the receptor as previously discussed^{12,14}. That C3 does bind directly to intact cells has been documented previously for lymphocytes¹⁷, but has not been evaluated for macrophages. Based on these data there may be a parallel between the IgG–Fc γ receptor interaction and the interaction of C3 and C3b with their receptor. Although both C3 and IgG can bind to their respective receptors, they exhibit a higher net affinity of binding on modification: by fixation to antigen allowing multi-point attachment in the case of IgG, or by cleavage and removal of the C3a fragment in the case of C3. In addition, receptor-bearing effector cells may interact more efficiently with membrane-bound C3b molecules clustered about an antigenic site than C3 molecules in solution.

Previous attempts to purify the C3b receptor from solubilized cells have been unsuccessful^{11,18}. Recently, using affinity chromatography with C3-Sepharose, Fearon¹⁹ isolated a 205,000-MW protein from human erythrocytes which has been implicated as the C3b receptor²⁰, although ligand-binding activity of the purified material was not reported. Its relationship to the complement-binding macrophage glycoprotein described here remains to be defined.

We acknowledge the technical assistance of V. Krause and thank Drs C. Hammer and M. Frank for providing purified C3 for use in preliminary experiments.

Received 2 July 1980; accepted 2 February 1981.

1. Bianco, C. in *Biological Application Systems in Immunology* (eds Day, N. K. & Good, R. A.) 69–84 (Plenum Medical Book Co., New York, 1977).
2. Gelfand, M. C., Frank, M. M. & Green, I. *J. exp. Med.* **142**, 1029–1034 (1975).
3. Tack, B. F. & Prahl, J. W. *Biochemistry* **15**, 4513–4521 (1976).
4. Law, S. K., Lichtenberg, N. A. & Levine, R. P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7194–7198 (1980).
5. Bokisch, V. A., Muller-Eberhard, H. J. & Cochrane, C. G. *J. exp. Med.* **129**, 1027–1048 (1969).

6. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
7. Kulczycki, A. Jr, Krause, V., Chew Killion, C. & Atkinson, J. P. *J. Immun.* **124**, 2772–2779 (1980).
8. Cuatrecasas, P. *J. biol. Chem.* **245**, 3059–3065 (1970).
9. Kulczycki, A. Jr & Parker, C. W. *J. biol. Chem.* **254**, 3187–3193 (1979).
10. Kulczycki, A. Jr, Krause, V., Chew Killion, C. & Atkinson, J. P. *J. immun. Meth.* **37**, 133–138 (1980).
11. Pepys, M. B., Bell, A. J. & Rowe, I. R. *Scand. J. Immun.* **4**, Suppl. 3, 75–78 (1975).
12. Kulczycki, A. Jr, Hempstead, B. L., Hofmann, S. L., Wood, E. W. & Parker, C. W. *J. biol. Chem.* **254**, 3194–3200 (1979).
13. Schneider, R. J., Atkinson, J. P., Krause, V. & Kulczycki, A. Jr *J. Immun.* **126**, 735–740 (1981).
14. Hempstead, B. L., Parker, C. W. & Kulczycki, A. Jr *J. Immun.* **123**, 2283–2291 (1979).
15. Ishizaka, T., Ishizaka, K., Conrad, D. H. & Froese, A. J. *Allergy clin. Immun.* **61**, 320–330 (1978).
16. Isersky, C., Taurog, J. D., Poy, G. & Metzger, H. *J. Immun.* **121**, 549–558 (1978).
17. Theofilopoulos, A. N., Bokisch, V. A. & Dixon, F. J. *J. exp. Med.* **139**, 696–711 (1974).
18. Dierich, M. P. & Reisfeld, R. A. *J. Immun.* **114**, 1676–1682 (1975).
19. Fearon, D. T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5867–5871 (1979).
20. Fearon, D. T. *J. exp. Med.* **152**, 20–30 (1980).
21. Davis, B. J. *Ann. N. Y. Acad. Sci.* **121**, 404–427 (1964).

Cyclosporin A blocks receptors for HLA-DR antigens on T cells

Ronald Palacios* & Göran Möller

Department of Immunobiology, Karolinska Institute, Wallenberg Laboratory, Lilla Frescati, S-104 05, Stockholm, Sweden

Recent studies have shown that cyclosporin A, an undecapeptide extracted from two species of fungi (*Cylindrocarpum lucidum* and *Tricoderma polysporum*), has potent immunosuppressive effects with low myelocytotoxicity¹. In different experimental systems in both animals and humans, cyclosporin A administration decreased antibody production to T cell-dependent antigens, prolonged survival of skin, kidney and heart allografts^{1–3}, inhibited cytotoxic activity generated in the allogeneic mixed lymphocyte reaction⁴ and abrogated suppressor T-cell activity induced by concanavalin A (ref. 5). However, the mode of action of cyclosporin A and its role in establishing tolerance remains unclear. Continuous proliferation of T cells after stimulation with antigens or mitogens is maintained by growth factors, released by lectin- or antigen-activated T cells^{5,6}. One of these factors, interleukin 2 (IL-2), supports growth of activated but not of resting T cells^{5–6}, so that, to achieve a proliferative response a cell population must be 'activated' both to produce IL-2 and to become responsive to it. In the autologous mixed lymphocyte reaction (AMLR), purified human T lymphocytes proliferate when co-cultured with autologous B cells and/or macrophages^{7,8}. Recently, we found that antibodies against either the heavy non-polymorphic chain or the light polymorphic chain of HLA-DR antigens of the stimulator cells, prevented T cells from acquiring responsiveness to IL-2, inhibited the production of the growth factor and abrogated the T-cell proliferative response in AMLR and hapten-labelled autologous cell systems^{9,10}. These results suggested that the HLA-DR antigens rendered resting T cells sensitive to IL-2 and participated actively in the production of the growth factor. We have now studied the effect of cyclosporin A on the AMLR response, and found that the drug inhibited the proliferative response by blocking the receptors for HLA-DR antigens on T cells.

The addition of as little as 1 $\mu\text{g ml}^{-1}$ cyclosporin A to the AMLR cultures strongly inhibited the proliferative response when compared with control cultures. This inhibition occurred very early in the process of activation of the responding cells, for it was not seen if the drug was added 72 h after the initiation of the cultures (Fig. 1). Pretreatment of the T cells with 1 $\mu\text{g ml}^{-1}$ or more cyclosporin A abrogated proliferation in the AMLR, whereas treatment of the stimulator cells with up to 5 $\mu\text{g ml}^{-1}$ had no effect on the AMLR response. This indicates that cyclosporin A exerts its inhibitory effects by acting on the responding T cells rather than on the non-T stimulator cells, in

* To whom all correspondence should be sent.

agreement with previous studies which postulated a selective activity of cyclosporin A on T cells^{2,4}.

To define the mechanism by which cyclosporin A inhibits the AMLR response, we studied the sensitivity of the responding T cells to IL-2 in the presence or absence of the drug and its effect on the synthesis of IL-2 in the AMLR. The results show that T cells from AMLR cultures containing cyclosporin A did not respond to IL-2 stimulation, whereas T cells from cultures where no drug was added proliferated under the stimulus of IL-2 (Table 1). When previously activated T cells (by AMLR or by phytohaemagglutinin, PHA) were stimulated with different concentrations of IL-2 in the presence or absence of cyclosporin A inhibition did not occur, indicating that once T cells have responded to IL-2, they are no longer sensitive to the inhibitory effects of cyclosporin A (data not shown). These results, together with our previous findings that cyclosporin A-treated T cells are unable to absorb IL-2 whereas activated T cells are able to do so⁵, support the notion that the drug inhibits expression of IL-2 receptors on T cells^{5,11}. The production of IL-2 in the AMLR was also inhibited by the addition of cyclosporin A to the cultures (Table 2), suggesting a second mechanism by which this drug could inhibit T-cell activation in the AMLR.

It seems unlikely that cyclosporin A exerts its inhibitory effects by nonspecific toxic or lytic effects, for at the doses used in this study ($1-5 \mu\text{g ml}^{-1}$) the viability of the cells as determined by Trypan blue exclusion from AMLR cultures in the presence or absence of cyclosporin A were similar at the end of the culture period. Furthermore, the experiments carried out to determine

Table 1 Response to IL-2 of T cells from AMLR cultures performed in the presence or absence of cyclosporin A

Cells	Dilutions of IL-2		
	1:2	1:4	1:8
T cells cultured alone	$1.1 \pm 0.1^*$	1.0 ± 0.3	0.9 ± 0.1
T AMLR	9.1 ± 1.3	6.5 ± 0.2	4.1 ± 0.5
T AMLR + $1 \mu\text{g ml}^{-1}$ cyclosporin A	1.3 ± 0.4	0.8 ± 0.4	0.8 ± 0.2
T AMLR + $3 \mu\text{g ml}^{-1}$ cyclosporin A	1.0 ± 0.3	0.7 ± 0.2	0.9 ± 0.1

Cell separation: Peripheral blood mononuclear cells (MNC) from healthy adult volunteers were obtained as previously described^{5,7}. T and non-T cells were isolated by rosetting MNC with sheep red blood cells (SRBC) as described elsewhere⁷. After osmotic lysis of the erythrocytes, the lymphocytes were washed and resuspended in RPMI 1640 medium (Gibco) supplemented with 10% of heat-inactivated human AB serum, $20 \mu\text{g ml}^{-1}$ of gentamicin, glutamine and 1% of HEPES buffer solution (Gibco) which will be referred to as complete culture medium (CCM). Autologous mixed lymphocyte reaction cultures: AMLR cultures were performed as previously described⁷. Briefly, 10^5 T cells were cultured with 10^5 irradiated autologous non-T cells in round-bottom microtitre plates (Nunc) Delta, Nunc). The secondary AMLR cultures were established by co-culturing 10^5 T cells from each primary AMLR culture with 10^5 of the following irradiated ($2,500 \text{ rad}$) autologous stimulator cells: non-T cells, T cells activated by PHA for 8 days (T-PHA), T cells from AMLR cultures performed in the presence or absence of cyclosporin A and peripheral blood MNC stimulated with PHA for 4 days (MNC-PHA). Cultures in triplicate were incubated at 37°C in a humidified atmosphere of 5% CO_2 in air. $0.5 \mu\text{Ci}$ of ^3H -thymidine (Radiochemical Centre) was added to each well 12 h before the end of a 7-day culture for primary AMLR cultures and a 3-day culture period for secondary cultures. ^3H -thymidine uptake of the cultures was determined as previously described^{6,7}. Cyclosporin A was tested for its ability to inhibit the proliferative response in AMLR by adding different doses of the drug ($1-5 \mu\text{g ml}^{-1}$) to the cultures either at the initiation of or at different intervals during culture. In no instance was the concentration of ethanol (in which cyclosporin A was dissolved) greater than 0.25% whether in the cultures containing this drug or in medium (control). IL-2 obtained from PHA-stimulated MNC was added to 10^2 T cells of the following groups: T cells cultured but not stimulated, T AMLR = T cells from AMLR cultures, T AMLR + $1 \mu\text{g}$ or $3 \mu\text{g}$ cyclosporin A = T cells from AMLR cultures treated with cyclosporin A. Proliferation was determined by the incorporation of ^3H -thymidine during the last 12 h of the 2-day culture period.

* c.p.m. $\times 10^{-3}$, mean $\pm$ s.d. of seven independent experiments.

Table 2 Effect of cyclosporin A on the synthesis of IL-2 in the AMLR

Supernatant source	Dilutions of IL-2		
	1:2	1:4	1:8
T cells culture alone	$1.1 \pm 0.2^*$	1.0 ± 0.1	1.0 ± 0.2
AMLR	15.2 ± 2.4	10.6 ± 1.5	8.4 ± 0.9
AMLR + $1 \mu\text{g ml}^{-1}$ cyclosporin A	2.9 ± 0.7	1.9 ± 0.2	1.8 ± 0.4
AMLR + $3 \mu\text{g ml}^{-1}$ cyclosporin A	1.6 ± 0.5	1.8 ± 0.3	1.5 ± 0.2

IL-2 production. IL-2 was obtained from AMLR cultures as previously described⁵. Briefly, 3×10^6 T cells and 3×10^6 irradiated autologous non-T cells suspended in 1 ml of CCM in the presence or absence of 1 and $3 \mu\text{g ml}^{-1}$ of cyclosporin A were incubated in 17×100 -mm round-bottom plastic tubes (Falcon) at 37°C for 48 h. After the first 24 h of culture, the cells were washed three times in CCM, resuspended in CCM in the absence of cyclosporin A and incubated for the last 24 h of culture. After this, the tubes were centrifuged at 2,000 r.p.m. for 15 min, and the supernatants collected, filtered through $0.45 \mu\text{m}$ and stored at -20°C until used. The experiments performed to find out whether cyclosporin A inhibits the proliferative response of previously activated T cells to IL-2 stimulation were established as follows: 10^4 T cells, activated in AMLR or by PHA and maintained in culture with conditioned media containing IL-2 for at least 3 weeks, were placed in flat-bottom microtitre plates. Three different dilutions of AMLR supernatants with proved IL-2 activity or IL-2 obtained after 48 h of PHA stimulation of peripheral blood MNC as described by Bonnard *et al.*¹⁶ was added with cyclosporin A (1 and $3 \mu\text{g ml}^{-1}$) or medium to each well. Cultures in triplicate were incubated at 37°C for 48 h and proliferation during the last 12 h was determined by the incorporation of ^3H -thymidine as described elsewhere^{3,7}. Supernatants from the following cultures: T cells cultured but not stimulated, AMLR = T cells and irradiated autologous non-T cells, AMLR + $1 \mu\text{g}$ or $3 \mu\text{g ml}^{-1}$ cyclosporin A = T cells co-cultured with irradiated autologous non-T cells in the presence of cyclosporin A, were placed at three different final dilutions in each well of microtitre plates containing 10^4 T cells activated by PHA and maintained in culture with conditioned media plus IL-2 for 3 weeks. After 48 h of incubation, incorporation of ^3H -thymidine was determined (see Table 1 legend).

* c.p.m. $\times 10^{-3}$, mean $\pm$ s.d. of seven separate experiments.

the effects of this drug on the response to IL-2 stimulation of previously activated T cells revealed that cyclosporin A did not inhibit the proliferation of these cells after IL-2 stimulation. Moreover, the observation (Fig. 1) that T-cell proliferation was not inhibited when cyclosporin A was added 72 h after the initiation of AMLR cultures eliminates the possibility of nonspecific toxic or lytic effects on the responding cells. Finally, activation of the stimulator cells with Epstein-Barr virus was not affected by the presence of cyclosporin A in the cultures, and pretreatment of the stimulator cells with this drug did not modify their capacity to induce proliferation (data not shown).

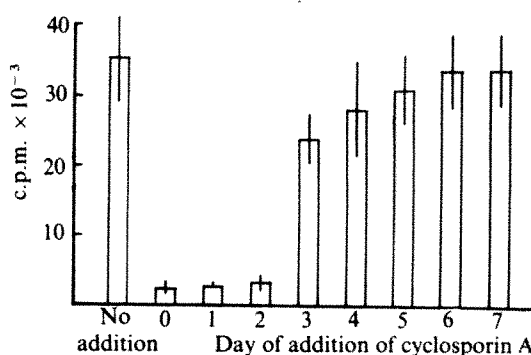


Fig. 1 Kinetic study of the inhibitory effect of cyclosporin A on the AMLR response. AMLR between 10^5 T cells and 10^5 irradiated autologous non-T cells were co-cultured in a final volume of 0.2 ml of complete culture medium (see Table 1 legend). Cyclosporin A ($1 \mu\text{g ml}^{-1}$) or medium was added to the cultures at different intervals after initiation of the AMLR cultures. Proliferation was determined by the ^3H -thymidine uptake during the last 12 h of the 7-day culture period.

Table 3 Response to various autologous stimulator cells of T cells from primary AMLR cultures performed in the presence or absence of cyclosporin A

Primary culture	Cyclosporin A addition ($\mu\text{g ml}^{-1}$)	Secondary AMLR culture stimulator cells				
		Non-T	T-PHA	T-AMLR	MNC-PHA	T-(AMLR + CYA)
T cells culture alone	—	$1.3 \pm 0.1^*$	1.8 ± 0.2	1.9 ± 0.1	2.4 ± 0.3	0.7 ± 0.1
	1	1.0 ± 0.1	0.7 ± 0.1	0.6 ± 0.2	0.9 ± 0.1	0.5 ± 0.1
T AMLR	—	33.4 ± 2.8	29.3 ± 1.7	20.6 ± 2.5	31.5 ± 1.9	2.9 ± 0.7
	1	6.3 ± 1.7	5.0 ± 1.8	4.8 ± 2.0	7.6 ± 2.3	1.1 ± 0.2
	3	4.2 ± 1.5	3.5 ± 0.6	3.3 ± 1.1	4.1 ± 1.6	1.4 ± 0.5
T AMLR + $1 \mu\text{g ml}^{-1}$ cyclosporin A	—	1.9 ± 0.1	1.3 ± 0.1	1.5 ± 0.3	4.2 ± 1.0	0.6 ± 0.1
T AMLR + $3 \mu\text{g ml}^{-1}$ cyclosporin A	—	1.1 ± 0.3	1.0 ± 0.1	1.4 ± 0.4	4.1 ± 0.7	0.9 ± 0.3

Cyclosporin A (1 or $3 \mu\text{g ml}^{-1}$) or medium was added to each well at the initiation of the secondary AMLR culture. T cells cultured alone or with irradiated autologous non-T cells in the presence (T-AMLR + 1 or $3 \mu\text{g ml}^{-1}$ cyclosporin A) or absence (T-AMLR) of cyclosporin A for 10 days (primary AMLR culture) were recovered by rosetting with SRBC and restimulated with the following autologous cells: Non-T cells; T-PHA = 8-day PHA-stimulated T cells; T-AMLR = T cells from 7-day cultures; MNC-PHA = 4-day PHA-stimulated peripheral blood MNC; T-AMLR = T cells from 7-day AMLR culture done in the presence of $1 \mu\text{g ml}^{-1}$ cyclosporin A. The cell mixture was incubated at 37°C for 3 days, and the ^3H -thymidine incorporation during the last 12 h of the culture period determined (see Table 1 legend).

* Mean $\pm$ s.d. of three independent experiments expressed as c.p.m. $\times 10^{-3}$.

We next studied the effects of cyclosporin A on the secondary AMLR using different kinds of irradiated autologous cells as stimulators. T cells activated in primary AMLR cultures or by PHA possess DR antigens on their surface^{9,12,13} and can function as stimulator cells in the AMLR⁹. Table 3 shows that the addition of cyclosporin A at the start of the secondary AMLR cultures inhibited the proliferation of the responding T cells. Interestingly, T cells from primary AMLR cultures performed in the presence of cyclosporin A were unable to function as stimulators in secondary AMLR cultures, whereas T cells from primary cultures without cyclosporin A functioned well as stimulator cells. Furthermore, the former cells were not killed by treatment with a monoclonal anti-DR antibody (NE011, NEN) plus complement whereas 50–60% of the latter cells were lysed by similar treatment. This indicates that T cells exposed to cyclosporin A lose their capacity to bind DR antigens, a property of T cells which has been demonstrated by Elliot *et al.*¹⁴

We also found that T cells from primary cultures containing cyclosporin A exhibited a proliferative activity similar to that shown by the population of cultured, not stimulated, T cells, whereas non-cyclosporin A-treated T cells responded with a typical pattern of secondary responses (higher proliferation at an earlier time). These findings suggest that cyclosporin A-treated T cells were not activated in the first culture, as they proliferate in secondary cultures to a similar extent to T cells cultured but not activated.

Thus, cyclosporin A inhibits the AMLR response by acting selectively on the responding T cells before they become responsive to IL-2, after which the drug can no longer inhibit T-cell proliferation and/or growth. We therefore conclude that cyclosporin A exerts its inhibitory effects by interfering with the process which normally enables resting T cells to become sensitive to IL-2, a step which is induced by signals from HLA-DR antigens of the stimulator cells^{9,10}. As cyclosporin A has no effect on stimulator cells, it probably blocks the receptors for the DR molecules on T cells and consequently inhibits the expression of IL-2 receptors on these cells. This notion is supported by several pieces of evidence. First, cyclosporin A-treated T cells are unresponsive and unable to absorb IL-2 (ref. 5 and the present study). Second, T cells from primary AMLR cultures performed in the presence of cyclosporin A do not possess DR antigens and are unable to function as stimulator cells in the AMLR. Third, cyclosporin A does not affect other receptors present on T cells, as it did not reduce the number of T cells forming rosettes with sheep red blood cells or autologous erythrocytes (data not shown).

We have recently found that HLA-DR antigens participate in the process of IL-2 production in either the AMLR or allogeneic mixed lymphocyte reaction (MLR) by rendering IL-2-producer T cells sensitive to interleukin 1 (formerly lymphocyte activating factor), which induces these cells to secrete the growth factor. The addition of cyclosporin A to AMLR and MLR cultures

causes IL-2-producer T cells to become unresponsive to interleukin 1 and abrogates the synthesis of IL-2¹⁵.

This work was supported by grants from the Swedish Cancer Society and the Swedish MRC.

Received 13 October 1980; accepted 11 February 1981.

1. Borel, J. F., Feuer, C., Gubler, H. U. & Stahelin, H. Ag. *Actions* **6**, 468–475 (1976).
2. Borel, J. F. *Immunology* **31**, 631–641 (1976).
3. White, D. J. G., Calne, R. Y. & Plumb, A. *Transplant Proc.* **11**, 855–859 (1979).
4. Horsburgh, T., Wood, P. & Brent, L. *Nature* **286**, 609–611 (1980).
5. Palacios, R. & Möller, G. *J. exp. Med.* **153**, (1981).
6. Smith, K. A. *Immun. Rev.* **51**, 338–357 (1980).
7. Palacios, R., Llorente, L., Alarcón-Segovia, D., Ruiz-Arguelles, A. & Diaz-Jouanen, E. *J. clin. Invest.* **65**, 1527–1530 (1980).
8. Opelz, G., Takasugi, M. & Terasaki, P. I. *J. exp. Med.* **142**, 1327–1333 (1975).
9. Palacios, R. *et al.* (submitted).
10. Palacios, R., Möller, G., Claesson, L. & Peterson, P. A. (submitted).
11. Larsson, E. L. *J. Immun.* **124**, 2828–2833 (1980).
12. Elliot, B. E., Nagy, Z., Nabholz, M. & Pernis, B. *Eur. J. Immun.* **7**, 287–291 (1977).
13. Engleman, E. G., Benike, C. J. & Charron, D. J. *J. exp. Med.* **152**, Suppl., 114–126 (1980).
14. Elliot, B. E., Nagy, Z. A., Ben-Neriah, Y. & Givol, D. *Nature* **285**, 496–498 (1980).
15. Palacios, R. (submitted).
16. Bonnard, G. D., Yasaka, K. & Maca, R. D. *Cell. Immun.* **51**, 390–401 (1980).

Role of RNA transcripts in replication incompatibility and copy number control in antibiotic resistance plasmid derivatives

Jonathan Rosen*, Thomas Ryder, Hisako Ohtsubo & Eiichi Ohtsubo

Department of Microbiology, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

The genes required for autonomous replication and incompatibility in the antibiotic resistance plasmids R100 and R1 have been located within a 2.5-kilobase region of the 90-kilobase genome, within which the incompatibility gene occupies a 1.3-kilobase region excluding the replication origin. We now report that three RNA species are synthesized *in vitro* from the 2.5-kilobase region, which R100 and R1 have in common. One, a long RNA molecule which is transcribed in the direction of DNA replication, probably acts as a messenger or a protein required for plasmid replication. The second RNA species, only 91 nucleotides long, is transcribed in the opposite direction, from a region of the DNA entirely contained within the first and known to specify incompatibility and copy control functions. The third RNA species, 150 bases long, is transcribed from a region including the replication origin¹; it may be a primer of DNA synthesis or, in conjunction with the second of the three RNA species, an influence in the control of replication.

When the *Pst*I fragment which carries the origin of replication of pSM1 or pTR1 is transcribed *in vitro* in the presence of

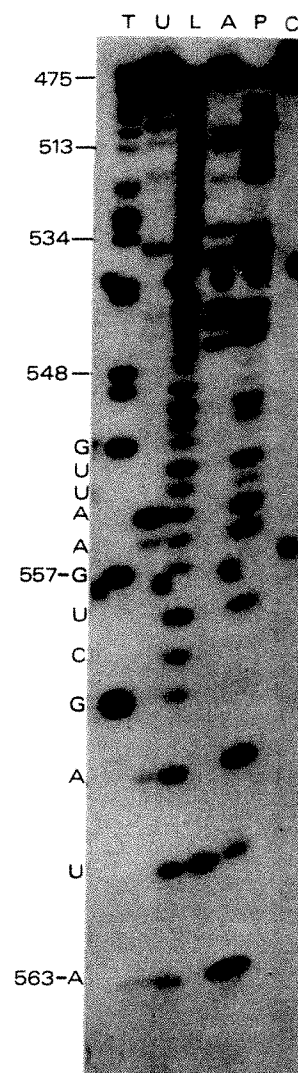


Fig. 2 Autoradiogram of a sequencing gel of the 5' end of RNAI. The position of selected G residues on the sequence shown in Fig. 3 are marked. Sequencing was performed by enzymatic cleavage as described by Donniss-Keller *et al.*¹⁵ and Simoncsits *et al.*¹⁶. The lanes of the gel are as follows: C, control; P, Phyl; A, pancreatic RNase A; L, alkaline digest; U, RNase U₂; T, RNase T₁. RNA was synthesized *in vitro*, basically as described by Majors¹⁷ in a mixture containing 0.2 pmol DNA (circular plasmids or linear restriction fragments) 30 mM Tris (pH 8.0), 2 mM MgCl₂, 0.1 M KCl, 0.1 mM EDTA, 0.1 mM dithiothreitol and 0.5 mg ml⁻¹ bovine serum albumin. RNA polymerase (Boehringer Mannheim) was added in a molar ratio of 10:1 enzyme to DNA and the mixture was incubated for 10 min at 37°C. Nucleotides were added to a final concentration of 20 µM except for the labelled nucleotide, which was added to 2 µM. The mixture was then transferred to a tube containing 40 µCi [α -³²P]UTP or 0.2 mCi dried down [γ -³²P]ATP¹⁸ or [γ -³²P]GTP (>1,000 Ci mmol⁻¹, ICN) and incubated at 37°C for 30 min. The transcripts were purified by gel electrophoresis as described by Donniss-Keller *et al.*¹⁵

[γ - ^{32}P]ATP, a third transcript called RNAIII is found. The 5' end of this transcript occurs at nucleotide 2,316, and the region upstream of this site displays moderate homology to prototypical promoter sites (Fig. 3). The transcript has two termination sites when this *Pst*I fragment is used as a template, the first being simply the end of the *Pst*I fragment, and the second at approximately nucleotide 2,465, about seven bases upstream of the *Pst*I cutting site. Label is preferentially incorporated in the shorter transcript with a ratio of 3.7:1. The sequence in this region, TAATCAATAT, is remarkably similar to that at the termination site of λtR1 -CAATCAA⁴-and tRNA_i^{Tyr}-CAATCAATAT⁵-both of which are ρ -factor-dependent transcriptional termination sites.

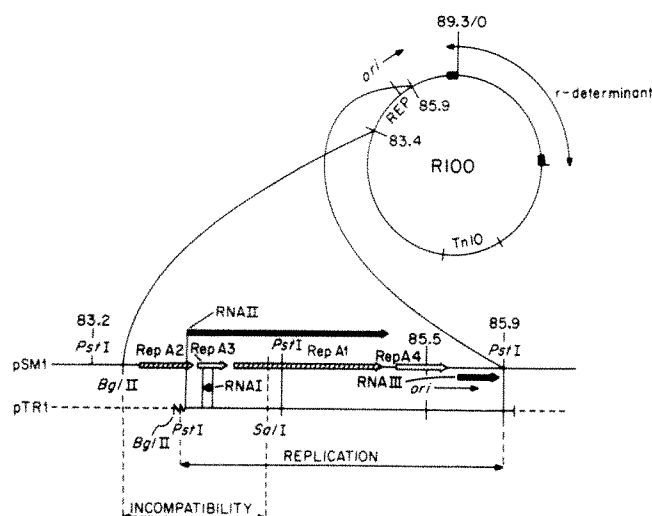


Fig. 1 A circular representation of some genetic and physical properties of the plasmid R100. The region required for replication (REP) includes the origin of replication (*ori*) from which replication proceeds unidirectionally in the direction shown. The origin region was determined by electron microscopic analysis of replicating molecules¹⁴ and the sequence was determined¹. The figure also shows a linear representation of the replication regions of the R100 derivative pSM1 and the R1 derivative pTR1 which was constructed from Rsc13 such that only the replication region and the gene conferring ampicillin resistance remain (T.R., unpublished results). The regions required for replication and incompatibility are shown at the bottom of the figure. The open arrows marked RepA3 and RepA4 indicate open reading frames predicted by nucleotide sequence analysis². The cross-hatched arrows show the position of reading frames whose polypeptide products have been confirmed². The heavy arrows marked RNAI, RNAII and RNAPIII are the transcripts described in the text.

* Present address: Department of Pathology, Stanford University School of Medicine, Stanford, California 94305, USA.

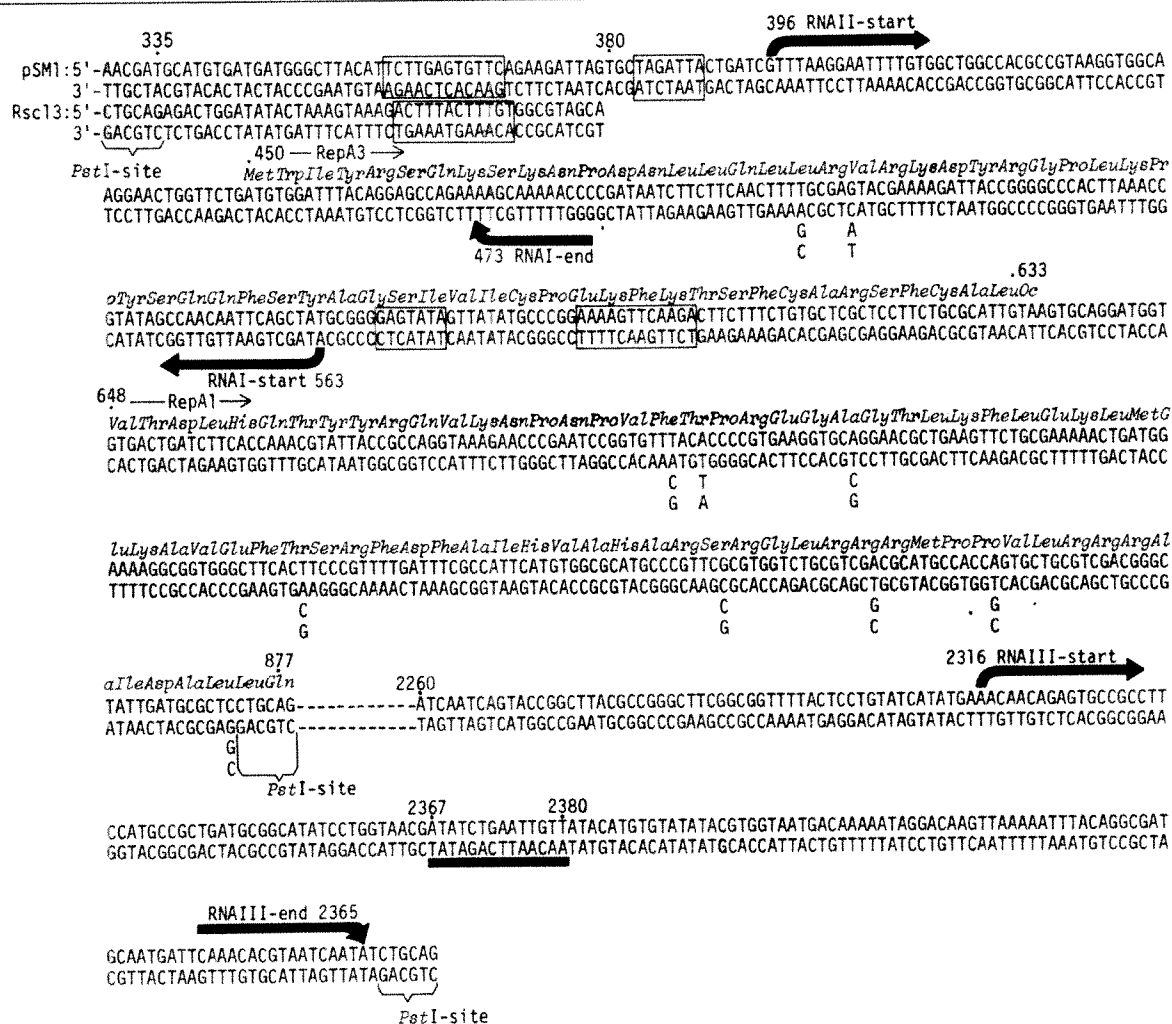


Fig. 3 The location of RNAI and RNAII within a *Pst*I fragment of pTR1. The entire sequence of the plasmid pSM1 within this fragment is shown and the differences between pSM1 and pTR1 are shown below the pSM1 sequence. The heavy lines above and below the sequence show the location of RNAI and RNAII at the nucleotide sequence level. The boxed regions proximal to each transcript delineate the Pribnow and -35 recognition sequences. The coding regions for RepA1 and RepA3 as well as the amino acid residues encoded are indicated on the sequence. A portion of the sequence of the origin region of pSM1 (2,260-2,480) including RNAIII and the sequence complementary to the 5' end of RNAI is also shown.

RNAIII has only been found when the linear *Pst*I fragment is transcribed, and its absence when superhelical pSM1 or pTR1 DNA is used as template may be due to nonspecific termination in the absence of ρ -factor yielding a heterogeneous range of transcripts. Most interestingly, the region of complementarity between RNAI and the origin region occurs within this transcript (Fig. 3). As will be discussed below, the interaction between RNAI and the superhelical template may affect the transcription of RNAIII.

RNAII is almost certainly the messenger RNA required to translate a 33,000-molecular weight (MW) protein called RepA1 which is assumed to be necessary for replication initiation (see Fig. 1 and ref. 2). As RNAII spans the coding region of a second, much smaller hypothetical 7,000-MW protein called RepA3, it may also direct the synthesis of that protein. However, RepA3 has not been seen *in vivo* or *in vitro*, so it is unclear whether this hypothetical protein is translated.

An interesting hypothesis is that RNAI may be partly responsible for copy number control and incompatibility, because it is synthesized in the region responsible for these functions. Figures 3 and 4 show that the two sequence differences between pSM1 and pTR1 in this region occur at the top of the proposed secondary structure for RNAI. Comparison of these sequences with the wild-type R1 DNA sequence recently obtained by Stougaard *et al.*⁶ reveals a single base-pair change between wild type and pSM1 at position 505 and

between wild type and pTR1 at position 510; these two differences between the pSM1 and pTR1 DNA sequences are probably responsible for their mutant copy number phenotypes, possibly altering the copy number and incompatibility properties of these plasmids by altering the affinity that RNAI has to a protein also involved in replication control.

RNAIII is likely to act either as the actual primer for the initiation of the plasmid DNA synthesis, or as a transcriptional activation signal which we suggest alters the conformation of a secondary structure at the replication origin upstream of this RNA transcript much as proposed by Hobom *et al.* for λ replication⁷. A large secondary structure has in fact been found upstream of RNAIII in pSM1 (ref. 1). As described above, 13 of the first 14 nucleotides of RNAI are complementary to the region between 2,367 and 2,380 of the replication origin region spanned by RNAIII (Figs 3, 4). Assuming that RNAI is involved in the incompatibility reaction and acts directly at the origin, RNAI may inhibit the initiation of DNA synthesis by binding to the complementary sequence at the origin region, probably with the aid of a protein(s).

In our *in vitro* system, RNAI and RNAII are transcribed simultaneously from the same region of the plasmid but in opposite directions. *In vivo*, the production of RNAI itself may be controlled at the transcriptional level by an attenuator-like mechanism^{8,9} affecting both copy number and incompatibility. If a ribosome succeeds in binding to a nascent RNAI before the

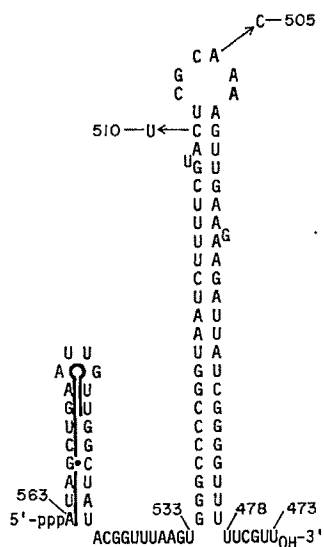


Fig. 4 Possible secondary structure of RNAI and homology to the origin of pSM1. The two arrows at the top of the loop indicate the differences between the sequence of RNAI and pSM1 and pTR1 at positions 505 (A→C) and 510 (C→U). The line within the first small hairpin loop indicates the region complementary to the origin region between nucleotides 2,367 and 2,380 of the pSM1 replication region sequence (see Fig. 3 and ref. 1).

RNA polymerase reaches the termination point the secondary structure would be disrupted and, analogously to the pattern of transcription in attenuation systems^{8,9}, transcription might proceed beyond the termination point reported here. It is difficult to predict if RNAI is actually translated. A small open reading frame begins at the start codon GUG (535–533) just before the large stem structure seen in Fig. 4 and continues until the termination codon UAG (384–382) (see Fig. 3). This reading frame is preceded by the sequence GGU (542–540), which may serve as a ribosome binding site¹⁰. The protein which could be translated from the extended RNA would have a MW of ~5,300, and as this is the region encoding copy number control and incompatibility functions, both of which are probably controlled by a repressor from this region, the 5,300-MW polypeptide is a possible candidate. However, because the base-pair change between pSM1 and R1 would not cause an amino acid change if the extended RNAI were translated, this small polypeptide would be identical in the wild-type and mutant plasmids. The two modes of transcription of RNAI may, however, correspond to the repressed and derepressed states of replication. Thus, the rate of plasmid replication could be tied, deterministically, to other aspects of cellular metabolism by the average rate of initiation for RNAI translation. This larger RNA species may no longer be effective either because of steric problems in binding at the origin or in forming a duplex with nascent RNAII or because of improper or altered processing of the molecule.

Recent data obtained in the ColE1 plasmid system shows a very similar pattern of transcription in the region required for replication to that seen in pSM1 and pTR1 (ref. 11), although the ColE1 plasmid does not require any of its own proteins for replication and has no transcript corresponding to RNAIII. A small RNA which may be analogous to RNAI is produced in the direction opposite to the direction of replication and ~450 bases upstream from the replication origin¹². A small insertion mutation that maps in this RNA increases the copy number of the plasmid, indicating that this RNA is an important control element¹³. Furthermore, it is transcribed in the direction opposite that of a second, larger transcript required *in vitro* as a primer for DNA synthesis¹¹. We do not know whether RNAII in our system can be used as both an mRNA and a primer for R-factor replication. It is reasonable to propose that RNAI, by interacting at the origin region, is involved in copy number control,

based on its location in the replication region, the changes found in pSM1 and pTR1 that appear to cause the increased copy number phenotype and by partial analogy to the ColE1 system. Further genetic and biochemical analyses to determine the nature of this control system are under way.

We thank Chiyoko Machida for sharing her expertise in the use of various RNases, Bert Semler and Naomi and Akiko Kitimura for help in the initial stages of RNA sequencing, Yvonne Tong and Linda Hollman for typing the manuscript, Joyce Schirmer for the illustrations and Jeff Demian for the photographs. This work was supported by the USPHS with research grants to E.O. (GM22007) and H.O. (GM26779) and by partial support to T.R. under training grant (CA-09176).

Note added in proof: In recent attempts to shorten the replication region we have been unable to delete the region coding for RNAI and RNAII but the region coding for RNAIII can be eliminated. Thus RNAIII is unlikely to be a primer for DNA synthesis but the possibility that its normal function includes some other aspect of replication control is not excluded.

Received 23 September 1980; accepted 28 January 1981.

1. Rosen J., Ohtsubo, H. & Ohtsubo, E. *Molec. gen. Genet.* **171**, 287-293 (1979).
2. Rosen, J., Ryder, T., Inokuchi, H., Ohtsubo, H. & Ohtsubo, E. *Molec. gen. Genet.* **179**, 527-537 (1980).
3. Rosenberg, M., Court, D. A. *Rev. Genet.* **13**, 319-353 (1979).
4. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. *Nature* **272**, 414-423 (1978).
5. Küpper, H., Sekiya, T., Rosenberg, M., Egan, J. & Landy, A. *Nature* **272**, 423-428 (1978).
6. Stougaard, P., Molin, S., Nordström, K. & Hansen, F. *Molec. gen. Genet.* (in the press).
7. Hobom, G. et al. *Cold Spring Harb. Symp. quant. Biol.* **43**, 165-178 (1978).
8. Lee, F. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4365-4369 (1977).
9. Keller, E. B. & Calvo, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6186-6190 (1979).
10. Shine, J. & Dalagarno, L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1342-1346 (1974).
11. Itoh, T. & Tomizawa, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2450-2454 (1980).
12. Levine, A. D. & Rupp, W. D. in *Microbiology 1978* (ed. Schlessinger, D.) (American Society for Microbiology, Washington DC, 1978).
13. Conrad, S. E. & Campbell, J. L. *Cell* **18**, 61-71 (1979).
14. Ohtsubo, E., Feingold, J., Ohtsubo, H., Mickel, S. & Bauer, W. *Plasmid* **1**, 8-18 (1977).
15. Donnissitts, H., Maxam, A. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527-2538 (1977).
16. Simoncsits, A., Brownlee, G. G., Brown, R. S., Robin, J. R. & Guillely, H. *Nature* **269**, 833-836 (1977).
17. Majors, J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4394-4398 (1975).
18. Johnson, R. A. & Walseth, T. F. *Adv. Cyclic Nucleotide Res.* **10**, 135-167 (1979).

An inducible DNA replication-cell division coupling mechanism in *E. coli*

Olivier Huisman & Richard D'Ari

Institut de Recherche en Biologie Moléculaire, 2, Place Jussieu
(Tour 43), 75221 Paris Cedex 05, France

Cell division is a tightly regulated periodic process. In steady-state cultures of Enterobacteriaceae, division takes place at a well defined cell mass¹ and is strictly coordinated with DNA replication². In wild-type *Escherichia coli* the formation of cells lacking DNA is very rare³, and interruptions of DNA replication arrest cell division⁴⁻⁶. The molecular bases of this replication-division coupling have been elusive but several models have been proposed. It has been suggested, for example, that the termination of a round of DNA replication may trigger a key event required for cell division⁴⁻⁸. A quite different model postulates the existence of a division inhibitor which prevents untimely division and whose synthesis is induced to high levels when DNA replication is perturbed⁹. The work reported here establishes the existence of the latter type of replication-division coupling in *E. coli*, and shows that the *sfiA* gene product is an inducible component of this division inhibition mechanism which is synthesized at high levels after perturbations of DNA replication.

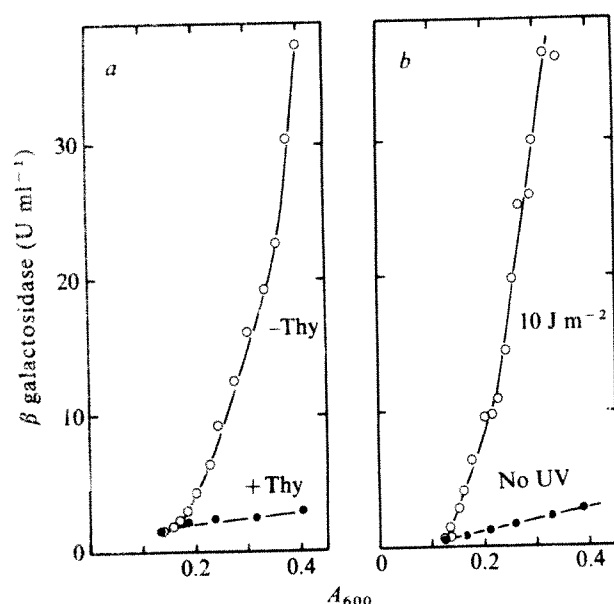


Fig. 1 Inducibility of *sfiA* gene expression by thymine starvation and UV-ray irradiation. Bacteria were grown at 37°C in M63 minimal salts medium³⁵ supplemented with glucose, casamino acids and growth requirements. At an absorbance *A* of ~0.4, the cultures were adjusted to *A* 0.1, a part of each was induced (a, centrifuged and washed twice, then resuspended in medium without thymine; b, irradiated with 10 J m⁻² UV ray) and incubation was continued at 37°C. Samples were withdrawn every 5 (induced cultures) or every 15 (uninduced cultures) min for absorbance measurements and β -galactosidase assays (carried out as in ref. 35). Enzyme units were calculated as described in ref. 37. Strain construction is described in Table 1 legend.

Much evidence points to the induction by DNA damaging treatments of a number of biological phenomena, collectively termed the SOS response (reviewed in refs 10–12). These phenomena include the appearance of new DNA repair and mutagenic activities, massive synthesis of the *recA* protein, lytic development of certain prophages in lysogenic strains and division inhibition (filamentous growth) in non-lysogens. The *recA* and *lexA* functions regulate the expression of all these phenomena. Mutant alleles of either locus can entirely prevent induction of the SOS response (*recA*⁻, *lexA*⁻) or cause its gratuitous expression at 42°C (*tif* at the *recA* locus, *tsl* at the *lexA* locus)^{10–12}.

Incubation of non-lysogenic *tif* or *tsl* strains at 42°C arrests cell division, resulting in the formation of long, multinucleate filaments. This filamentous growth is completely suppressed by *sfiA* mutations, whereas other aspects of the SOS response are unaffected^{13–15}. The *sfiA* locus thus defines an element directly involved in the division inhibition process.

One model of replication-division coupling predicts the existence of an inducible division inhibitor. We thus sought to determine whether synthesis of the *sfiA* product is induced when DNA replication is perturbed.

To obtain accurate quantitative measurements of the rates of *sfiA* gene expression in various conditions we chose the powerful method of operon fusion, using the phage Mu d(Ap, *lac*)¹⁶ to link the structural gene for β -galactosidase (*lacZ*) to the promoter of the *sfiA* operon. Near one extremity of the phage genome are the *lacZ* and *lacY* genes, lacking a promoter. The phage genome is so constructed that, when inserted in a host operon in the proper orientation, the *lac* genes of the prophage can be expressed from the adjacent bacterial promoter: their level of expression then reflects the level of expression of the bacterial operon in which the prophage is inserted.

Previous studies have shown that *sfi*-dependent filamentation occurs early during thymine starvation¹⁷. The effect of thymine starvation on *sfiA* expression was investigated in a *thyA sfiA::lacZ* fusion strain. The differential rate of β -galactosidase synthesis in these conditions quickly increases 20–40-fold (Fig. 1a, Table 1) reaching ~0.3% of total protein synthesis. Thus the *sfi*-dependent filamentation normally observed during thymine starvation is correlated with a large increase in the rate of *sfiA* gene expression. A UV-ray dose of 10 J m⁻² similarly causes a rapid 20–40-fold increase in the differential rate of β -galactosidase synthesis (Fig. 1b, Table 1). Nalidixic acid, a specific inhibitor of DNA gyrase, and methyl methanesulphonate, a DNA alkylating agent, also induce high rates of *sfiA* gene expression (Table 1). A Mu d(Ap, *lac*) prophage inserted in the *gal* operon does not lead to UV-ray inducible β -galactosidase synthesis (Table 1), showing that the inducibility observed in *sfiA::lacZ* strains depends on the *sfiA* location of the prophage. We conclude that perturbations of DNA replication induce a high level of expression of the *sfiA* gene.

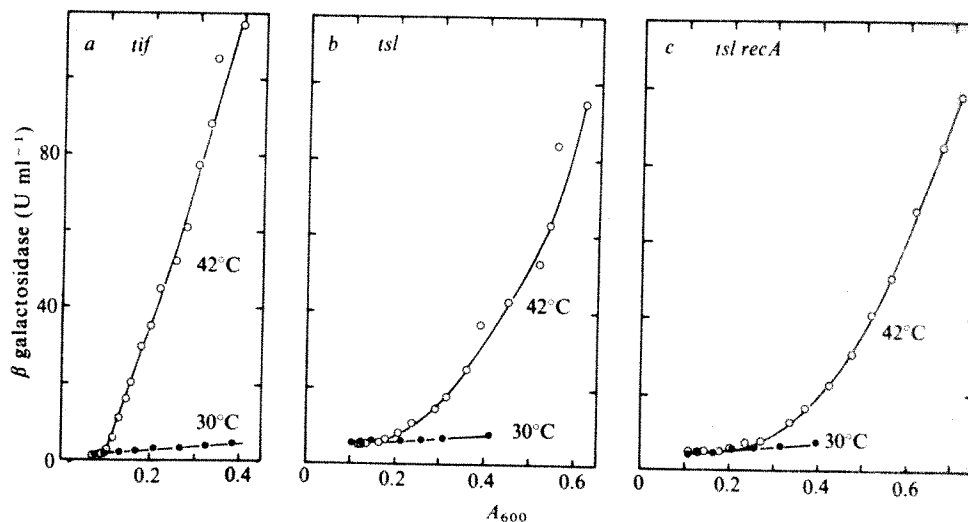
We next studied the effect on *sfiA* gene expression of mutations known to suppress or enhance *sfi*-dependent filamentation. *recA*⁻ and *lexA*⁻ mutations completely suppress the SOS

Table 1 Induction of *sfiA* gene expression

Strain	Inducing treatment	β -galactosidase specific activity (U per mg dry weight)		
		Un-induced	Induced	Dose
<i>sfiA::lacZ thyA</i>	Thymine starvation	50	410	
<i>sfiA::lacZ</i>	UV-ray irradiation	30	140	1 J m ⁻²
			440	10 J m ⁻²
<i>sfiA::lacZ</i>	Nalidixic acid	30	110	2 μ g ml ⁻¹
			710	40 μ g ml ⁻¹
<i>sfiA::lacZ</i>	Methyl methanesulphonate	30	360	0.025%
<i>sfiA::lacZ</i>	Nitrofurantoin	30	90	2 μ g ml ⁻¹
<i>sfiA::lacZ recA</i>	UV-ray irradiation	10	10	1 J m ⁻²
			10	10 J m ⁻²
<i>sfiA::lacZ lexA</i>	UV-ray irradiation	10	10	1 J m ⁻²
			10	10 J m ⁻²
<i>sfiA::lacZ tif</i>	42°C	80	700	
<i>sfiA::lacZ tsl</i>	42°C	200	450	
<i>sfiA::lacZ tsl recA</i>	42°C	200	400	
<i>sfiA::lacZ lon</i>	UV-ray irradiation	20	400	10 J m ⁻²
<i>gal::lacZ</i>	Galactose	10	50	0.4%
	UV-ray irradiation	10	10	10 J m ⁻²
<i>gal::lacZ recA</i>	Galactose + UV rays	10	25	0.4%, 10 J m ⁻²
<i>gal::lacZ lexA</i>	Galactose + UV rays	10	30	0.4%, 10 J m ⁻²
<i>gal::lacZ tif</i>	42°C	10	10	

Bacteria were grown at 37°C in M63 salts medium³⁵ supplemented with glucose, casamino acids and growth requirements (except for the experiments with *tif* and *tsl* strains, grown at 30°C) to a value of *A*₆₀₀ of ~0.4, then diluted to *A* 0.1 (the thymine-starved culture was first centrifuged and washed twice to remove exogenous thymine). One part of the culture was induced (by removal of thymine, UV-ray irradiation, temperature shift or addition of the appropriate factors), the cultures were further incubated for 60 min and β -galactosidase specific activities measured³⁵. Adenine (100 μ g ml⁻¹) was added to the *tif* cultures at the time of the temperature shift. To isolate Mu d(Ap, *lac*) insertions at the *sfiA* locus, we selected for a Sfi⁻ phenotype^{13,25,26} in a *lon* strain²⁴, taking advantage of the hypersensitivity of the strain to low concentrations of nitrofurantoin³⁶. Lysogenic Sfi⁻ clones were selected from a Mu d(Ap, *lac*)-infected *lon* culture on plates containing ampicillin (50 μ g ml⁻¹) and nitrofurantoin (2 μ g ml⁻¹). These were purified, checked for UV-ray resistance (to confirm the Sfi⁻ character) and analysed genetically. All Sfi⁻ clones carried a *sfiA* mutation, about 50% co-transducible with *pyrD* by P1^{13,15,25,26}. For most isolates the introduction by transduction of the Sfi⁻ character was invariably accompanied by loss of the Mu d(Ap, *lac*) prophage; the strains thus harboured a single Mu d(Ap, *lac*) prophage, located at the *sfiA* locus and conferring a Sfi⁻ phenotype. Strains of this type having a significant basal level of β -galactosidase were considered to have the *lac* genes fused to the *sfiA* promoter. One such fusion was transduced with P1 *vir* (selection Ap^R) into a strain carrying a chromosomal *lac* deletion (to prevent extraneous synthesis of β -galactosidase) and a Mu c⁺ prophage inserted in the *trp* operon (to prevent thermoinduction of the Mu d(Ap, *lac*) prophage, whose repressor is thermolabile). One Ap^R transductant, verified to have a Sfi⁻ phenotype, was used for this study; all other *sfiA::lacZ* fusion strains were derived from this transductant. The *gal::lacZ* strains were constructed in a similar way using a strain carrying Mu d(Ap, *lac*) inserted in the *galT* gene as donor in the transduction to Ap^R.

Fig. 2 Regulation of *sfiA* gene expression in *tif(recA)* and *tsl(lexA)* mutants. Bacteria were grown at 30°C (a, in M63 minimal salts medium³⁵ supplemented with glucose, casaminoacids, growth requirements and—at 42°C—adenine; b and c, in LB broth). At A 0.4, the cultures were diluted to A 0.1 in medium prewarmed to 42°C (open symbols) or 30°C (closed symbols) and incubation was continued. Samples were withdrawn every 5 (42°C cultures) or every 15 (30°C cultures) min for absorbance measurements and β -galactosidase assays^{35,37}. Strain construction is described in Table 1 legend.



response¹⁰⁻¹². In *recA* or *lexA* derivatives of the *sfiA::lacZ* fusion strain no UV-ray induction of β -galactosidase synthesis is observed in the dose range 1–10 J m⁻² (compare with Table 1). These same *recA* and *lexA* mutations introduced into a *gal::lacZ* fusion strain do not affect the inducibility of β -galactosidase synthesis by galactose, and the capacity of these highly UV-ray sensitive strains to synthesize β -galactosidase is not seriously reduced by 10 J m⁻² UV-ray irradiation (Table 1).

The *tif* and *tsl* mutations in the *recA* and *lexA* genes cause constitutive expression of some SOS functions at 42°C (refs 18, 19), including *sfi*-dependent filamentation^{13,14,17}. At this temperature in *tif sfiA::lacZ* (Fig. 2a, Table 1) and *tsl sfiA::lacZ* (Fig. 2b, Table 1) fusion strains, a high rate of β -galactosidase synthesis is induced, but no similar induction is observed in a *tif gal::lacZ* fusion strain (Table 1). Furthermore, division inhibition by expression of the *tsl* mutations, unlike all other known inducing treatments, does not depend on *recA*⁺ activity²⁰ and in a *tsl recA sfiA::lacZ* fusion strain a high level of β -galactosidase synthesis is still observed at 42°C (Fig. 2c, Table 1).

The *recA*⁺ and *lexA*⁺ gene products are thus regulators of *sfiA* gene expression, probably at the transcriptional level. The *recA* and *din* genes are similarly inducible in a *recA*⁺–*lexA*⁺–dependent manner^{19,21}, and in the case of the *recA* gene this regulation has been shown to be at the transcriptional level^{22,23}.

The *lon* mutation²⁴ amplifies *sfi*-dependent filamentation^{13,25,26}, sensitizing the cell to small doses of DNA damaging treatments. At least seven different enzymes involved in cell-wall biosynthesis are present at higher levels in *lon* strains than in *lon*⁺ strains²⁷, and in one case, that of the *gal* operon, the *lon* mutant has been reported to have increased transcription rates²⁸. The induction of *sfiA* gene expression by UV-ray irradiation was identical in *lon* and *lon*⁺ *sfiA::lacZ* fusion strains in basal level and induction kinetics (see Table 1). Thus *lon*⁺ function does not seem to be a transcriptional regulator of the *sfiA* operon. As *lon* mutants are also deficient in Deg protease activity²⁹, it is tempting to speculate that this protease is normally involved in eliminating excess *sfiA* protein, allowing rapid resumption of cell division as soon as normal DNA replication is restored.

Current models accounting for the regulation of the SOS response postulate that DNA damaging treatments result in the formation of a signal molecule by the DNA replication complex³⁰⁻³². This molecule activates the *recA* protein to a protease form which, in turn, cleaves certain cellular repressors, in particular the *cI* repressor of λ prophage³³ and the *lexA* protein³⁴. The results reported here show that a high rate of *sfiA* gene expression is part of the SOS response. The *sfiA* product is directly involved in SOS division inhibition and its synthesis is induced very rapidly during thymine starvation or after UV-ray

irradiation (Fig. 1). The *sfiA* product thus provides a highly sensitive link coupling cell division to DNA replication by the following probable sequence of events: perturbations or alterations of the DNA replication complex cause the generation of a signal molecule which activates the *recA* protein to its protease form. The *recA* protease then cleaves the *lexA* repressor, leading to derepression of the *sfiA* operon. Finally, the high concentration of *sfiA* protein thus produced results in an inhibition of cell division. In this sense the *sfiA* gene product can be considered an inducible division inhibitor, although its precise role in division inhibition remains to be elucidated.

We thank Maxime Schwartz for providing a *gal::lacZ* fusion strain and Micheline Jaques for helping in experimental work. This work was supported in part by grants ATP 4185 (CNRS-INRA) and ATP 72-79-104 (INSERM). This letter is dedicated to the memory of Jacqueline George, whose enthusiasm and insight initiated the research.

Received 4 November 1980; accepted 9 February 1981.

- Schaechter, M., Maaløe, O. & Kjeldgaard, N. O. *J. gen. Microbiol.* **19**, 592–606 (1958).
- Cooper, S. & Helmstetter, C. J. *molec. Biol.* **31**, 519–540 (1968).
- Howe, W. E. & Mount, D. W. *J. Bact.* **124**, 1113–1121 (1975).
- Helmstetter, C. E. & Pierucci, O. J. *Bact.* **95**, 1627–1633 (1968).
- Clark, D. J. *J. Bact.* **96**, 1214–1224 (1968).
- Donachie, W. D. *J. Bact.* **100**, 260–268 (1969).
- Jones, N. & Donachie, W. *Nature new Biol.* **243**, 100–103 (1973).
- Zaritsky, A. & Pritchard, R. *J. Bact.* **114**, 824–837 (1973).
- Witkin, E. M. *Proc. natn. Acad. Sci. U.S.A.* **57**, 1275–1279 (1967).
- Radman, M. in *Molecular Mechanisms of Repair of DNA* (eds Hanawalt, P. C. & Setlow, R. B.) 355–367 (Plenum, New York, 1975).
- Witkin, E. M. *Bact. Rev.* **40**, 869–907 (1976).
- Devoret, R. *Biochimie* **60**, 1135–1140 (1978).
- George, J., Castellazzi, M. & Buttin, G. *Molec. gen. Genet.* **140**, 309–332 (1975).
- McEntee, K. in *DNA Repair Mechanisms* Vol. 9 (eds Hanawalt, P. C., Friedberg, E. C. & Fox, C. F.) 349–359 (Academic, New York, 1978).
- Huisman, O., D'Ari, R. & George, J. *J. Bact.* **144**, 185–191 (1980).
- Casadaban, M. J. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4530–4533 (1979).
- Huisman, O., D'Ari, R. & George, J. *Molec. gen. Genet.* **177**, 629–636 (1980).
- Castellazzi, M., George, J. & Buttin, G. *Molec. gen. Genet.* **119**, 139–152 (1972).
- Gudas, L. J. *J. molec. Biol.* **104**, 567–587 (1976).
- Mount, D. W., Walker, A. C. & Kosel, C. J. *Bact.* **121**, 1203–1207 (1975).
- Kenyon, C. J. & Walker, G. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2819–2823 (1980).
- Gudas, L. J. & Pardee, A. B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2330–2334 (1975).
- McPartland, A., Green, L. & Echols, H. *Cell* **20**, 731–737 (1980).
- Howard-Flanders, P., Simson, E. & Theriot, L. *Genetics* **49**, 237–246 (1964).
- Gayda, R. C., Yamamoto, L. T. & Markovitz, A. J. *Bact.* **127**, 1208–1216 (1976).
- Johnson, B. F. *Genet. Res.* **30**, 273–286 (1977).
- Hua, S. & Markovitz, A. J. *Bact.* **110**, 1089–1099 (1972).
- Buchanan, C. E., Hua, S., Avni, H. & Markovitz, A. J. *Bact.* **114**, 891–893 (1973).
- Gottesman, S. & Zipser, D. J. *Bact.* **133**, 844–851 (1978).
- McEntee, K. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5275–5279 (1977).
- Gudas, L. J. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5280–5284 (1977).
- Emmerson, P. T. & West, S. C. *Molec. gen. Genet.* **155**, 77–85 (1977).
- Roberts, J. W., Roberts, C. W. & Craig, N. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4714–4718 (1978).
- Little, J. W., Edmiston, S. H., Pacelli, L. Z. & Mount, D. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3225–3229 (1980).
- Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor, New York, 1972).
- Kirby, E. P., Ruff, W. L. & Goldthwait, D. J. *Bact.* **111**, 447–453 (1972).
- Pardee, A. B., Jacob, F. & Monod, J. *J. molec. Biol.* **1**, 165–178 (1959).

Molecular cloning of foot and mouth disease virus genome and nucleotide sequences in the structural protein genes

John C. Boothroyd*, Peter E. Highfield*, George A. M. Cross*, David J. Rowlands†, Peter A. Lowe†, Fred Brown† & Timothy J. R. Harris†

* Department of Immunochemistry, Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK

† Department of Biochemistry, Animal Virus Research Institute, Pirbright, Woking, Surrey GU24 0NF, UK

Foot and mouth disease virus (FMDV), of the family Picornaviridae, consists of a single-stranded RNA (~8,000 nucleotides), the translation of which is initiated on the 3' side of a 150-nucleotide poly(C) tract and yields a single polypeptide which is processed by host cell proteases into four primary products (Fig. 1)¹⁻⁴. One or more virus-specified proteases further cleave these into the final products, the capsid proteins (VP1-4) being derived from the precursor p88 (for review see ref. 5). There are seven serotypes of the virus and as it has been shown that the immunizing activity of FMDV particles is associated primarily with VP1 (refs 6, 7), it seems likely that antigenic variation in FMDV is a result of changes in the structure of this protein. To further our understanding of this variation and as a first step in the possible development of FMDV vaccines from genetically manipulated microorganisms, we report here the construction and analysis of recombinant plasmids containing cDNA copies of the RNA. Comparison of the deduced amino acid sequence with the known polypeptide sequences shows that the NH₂-termini of VP2 and VP3 are conserved between the A and O serotypes whereas that of VP1 (the immunizing antigen) varies by as much as 42% between serotypes.

Using two strains of FMDV (serotypes A₁₀, strain 61 and O₁, strain BFS), double-stranded (ds) cDNA was synthesized⁸ and recombinant clones derived as described elsewhere⁹. Briefly, FMDV RNA (10 µg), primed with oligo (dT)₁₂₋₁₈, was used as a template for ds cDNA synthesis by reverse transcriptase. The ds cDNA so obtained ranged in size from 200 to 8,000 base pairs with a mean of ~1,000 base pairs as judged by agarose gel electrophoresis. The (dC)-tailed cDNA (0.05 pmol) was annealed with (dG)-tailed *Pst*I-digested pAT153 (0.2 pmol)¹⁰ and used to transform competent¹¹ *Escherichia coli* HB101 cells¹². Those clones derived from A₁₀ RNA are designated pFA10(61)/t whereas those from O₁ RNA are designated pFO1(BFS)/t; these are abbreviated here to pFA and pFO, respectively. In both the pFA and pFO clone banks ~90% of the tetracycline-resistant colonies were sensitive to ampicillin (100 µg ml⁻¹) which suggested that these colonies carried plasmids with ds cDNA inserted at the *Pst*I site. Further characterization was initially on the basis of size, using a modification of the single-colony lysis procedure¹³. The size distribution paralleled that of the ds cDNA used in the cloning (data not shown); the largest plasmid, pFA206, contained an insert of 5,800 base pairs.

Construction of restriction maps of the four largest plasmids from both sets of recombinants showed that, in the case of the pFA set, two plasmids, pFA76 and pFA206, contained overlapping ds cDNA inserts which together corresponded to ~7,200 base pairs (~90% of the FMDV genome). In the pFO set, the largest plasmid, pFO251, contained all the sequence (~4,800 base pairs) present in the other pFO plasmids analysed. The *Pst*I site was not regenerated at both ends of all of the inserts which were analysed, probably because of slight exonuclease degradation at some stage during the construction of the recombinants.

To align the restriction maps with the physical map of FMDV, ³²P-labelled fragments of type A₁₀ RNA produced by T₁ RNase digestion were used as probes. Two oligonucleotides, T₁-3 and T₁-4, were used initially as these are located 1,000 (±300) and 2,300 (±500) nucleotides respectively from the 5' end of the RNA¹⁴. Aliquots of a *Pst*I/*Ava*I double digest of pFA76 were resolved by agarose gel electrophoresis, transferred to nitrocellulose filters¹⁵ and probed with ³²P end-labelled oligonucleotides. The two probes hybridized to different fragments of the

Fig. 1 Organization of the FMDV genome. The relative positions of the genes on the FMDV RNA have been published elsewhere⁵. VP1, 2, 3 and 4 are the capsid proteins and are derived from the precursor p88. The ds cDNA inserts of four recombinant plasmids are aligned with this map on the basis of restriction endonuclease analysis, hybridization to ³²P-labelled T₁ oligonucleotides (see Fig. 2) and on DNA sequence analysis. The T₁ oligonucleotides 20 and 27 are positioned on the basis of previously published results^{14,19}, whereas oligonucleotides 3 and 4 have been located exactly as a result of the DNA sequence analysis. The regions of DNA sequence presented in

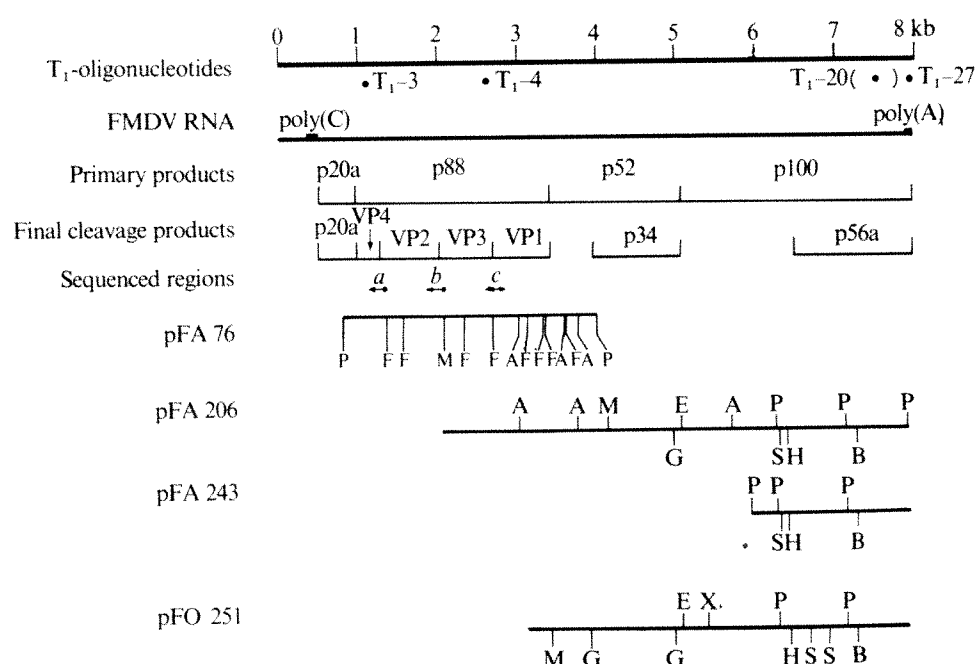


Fig. 3 are indicated by the heavy dots and arrowed lines over pFA76: a, b and c refer to the corresponding sections of Fig. 3. The restriction enzyme sites are designated: A, *Ava*I; B, *Bam*HI; E, *Eco*RI; G, *Bgl*II; H, *Hind*III; M, *Sma*I; P, *Pst*I; S, *Sal*I; X, *Xba*I; F, *Hinf*I. The *Ava*I and *Hinf*I maps have been completely determined for pFA76 only.

Table 1 Homology between the amino termini of VP1 of five FMDV strains

	A ₁₂	O _{1K}	O _{1B}	C _{3R}
A ₁₀	69	82	58 ± 4	71 ± 4
A ₁₂		65	73 ± 4	83 ± 4
O _{1K}			63 ± 2	67 ± 4
O _{1B}				73 ± 6

Percentage homologies were determined from the data presented in Fig. 3. The potential error accounts for the ambiguity in the published sequences.

digest (Fig. 2) and, as the relative positions of these restriction fragments in the pFA76 ds cDNA insert were known, the inserted cDNA could be aligned with the RNA to within 500 nucleotides and its orientation in the plasmid obtained. The results indicated that the pFA76 insert spanned from ~800 to 4,000 nucleotides from the 5' end of the genome and the insert in pFA206, as deduced by its overlap with pFA76, spanned from ~2,000 to 7,800 nucleotides from the 5' end. This alignment was consistent with the finding that T₁-20 (which had been mapped to between 500 and 1,000 nucleotides from the poly(A) tail¹⁴) specifically hybridized to the *Pst*I fragment of pFA206 nearest the 3' end (data not shown). The ds cDNA contained within these two recombinants therefore included the entire region coding for the major capsid proteins (Fig. 1).

The pFA recombinant clones were also screened by colony hybridization^{16,17} to detect plasmids containing inserts derived from the ends of the RNA. Short, oligo(dT)-primed ³²P-cDNA was used as a probe for plasmids containing sequences from the 3' end of the RNA; several colonies gave a positive signal with this probe. The sensitivity of this assay was increased by isolating DNA from 1-ml cultures of these recombinant bacteria¹⁸ and binding this DNA onto nitrocellulose filters. Oligonucleotide T₁-27, which is located 12–21 nucleotides from the 3' end of the RNA^{14,19}, was ³²P-labelled and used to probe such a filter. Only one recombinant clone (pFA243) gave a positive signal; by DNA sequence analysis, this recombinant was found to contain sequences corresponding to the 3' end of the virus RNA, including ~30 nucleotides of the poly(A) tail (data not shown). From a comparison of restriction maps, pFA243 and pFA206 overlapped by ~2,000 base pairs (Fig. 1).

To detect plasmids containing sequences from the 5' end of RNA ³²P-pCp-labelled S fragment, the small (~380 nucleotides²⁰), 5'-terminal fragment of the RNA produced by RNase H digestion in the presence of oligo(dG)₁₂₋₁₈ (ref. 3), was used to screen the recombinant clones. None of the clones gave a positive signal using this probe, which suggests that the 5' end of the RNA was not represented in the clone bank. In a similar hybridization with ³²P-labelled T₁-3 (located ~1,000 nucleotides from the 5' end, Fig. 1) only two recombinants (pFA76 and pFA187) gave positive signals (data not shown). Digests of pFA76 DNA were already known to hybridize with T₁-3 (Fig. 2, lane 3); further analysis of pFA187 by restriction endonuclease mapping showed that it extended only marginally further than pFA76 towards the 5' end (data not shown). These results indicate that ~750 nucleotides from the 5' end of the RNA are not represented in the recombinant clones.

An identical approach was used to orientate the pFO clones using the probes prepared from type A₁₀ RNA. This was possible even though there is only ~60% homology between the genomes of strains A₆₁ and O₁BFS²¹. Radiolabelled T₁-20 specifically hybridized to one fragment of a *Pst*I/*Eco*RI double digest of pFO251 (Fig. 2, lane 5). The orientation of the ds cDNA insert relative to the genome suggested by this result is supported by the similarity between the restriction maps of pFO251 and pFA206 (Fig. 1).

A detailed restriction map of pFA76 was constructed and used to sequence defined parts within the region coding for the structural proteins (Fig. 1). The sequences obtained around the junctions of these proteins are presented in Fig. 3 together with

the predicted amino acid sequences. The latter are compared with the partial amino acid sequences which have been published for these proteins derived from other strains and serotypes of FMDV²²⁻²⁴. The amino-terminal sequence of VP2 is identical in the A₁₀ and O_{1K} strains over the 20 amino acids available for comparison, while the amino terminus of VP3 shows just two differences out of 32 amino acids. The amino-terminal region of VP1 from A₁₀ can be compared with four other strains of FMDV (Fig. 3c). The relatively large extent of variation of the amino-terminal region of VP1 (compare with VP2 and VP3) may be related to the fact that this protein is the major immunizing antigen of FMDV^{6,7,25-27} and so variation of its sequence is under positive selection by the immune system of the host. The partial conservation of sequence which is apparent is probably due to constraints which must exist in order for VP1 to fulfill its function in the capsid and perhaps for the critical features of the protease cleavage site in p88 to be maintained. A comparison of the degree to which the sequences vary between strains (Table 1) shows no obvious correlation with the serological relationships of the different strains. However, as the region available for comparison constitutes only ~10% of VP1, it may be that the antigenic determinants which define its serotype lie elsewhere within the protein.

On the basis of their size and relative positions on the genome, it seems likely that the capsid proteins are contiguously coded²⁸.

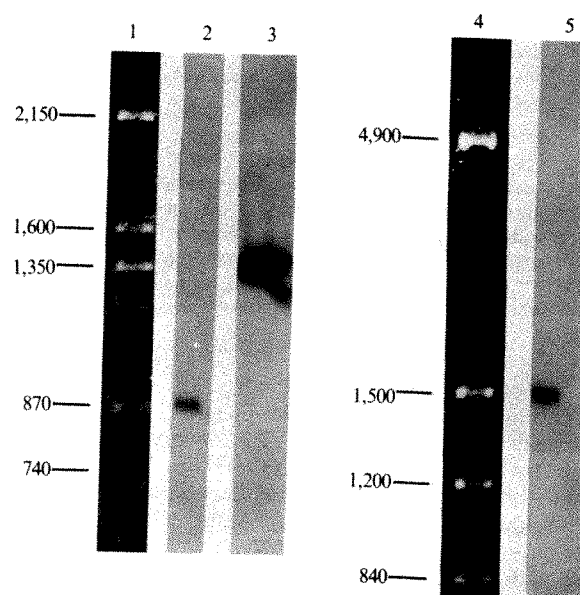


Fig. 2 Orientation of ds cDNA inserts with FMDV genome. Restriction fragments of pFA76 and pFO251 DNA were resolved by electrophoresis through a 0.85% agarose gel, and transferred to nitrocellulose filters (Schleicher & Schuell BA85)¹⁵. The filters were then incubated in 30 μ l cm⁻² of hybridization buffer (40% (v/v) formamide, 0.4 M NaCl, 10 mM PIPES-NaOH pH 6.4, 0.5% (w/v) SDS, 0.04% (w/v) bovine serum albumin, 0.04% (w/v) Ficoll, 0.04% (w/v) polyvinyl pyrrolidone 40 and 100 μ g ml⁻¹ *E. coli* tRNA) for 2 h at 50 °C before addition of the ³²P-labelled RNase T₁ oligonucleotide probes. The probes were prepared by two-dimensional gel electrophoresis of an RNase T₁ digest of 100 μ g of ³²P-labelled type A₁₀ RNA (5 $\times$ 10⁵ c.p.m.)³. The oligonucleotides were eluted from the gel and labelled at the 5' end using polynucleotide kinase as described elsewhere³⁰. The probes were allowed to hybridize for 16–20 h at 50 °C before the filters were washed in 2 $\times$ SSC (eight changes of 250 ml each) at 45 °C, air-dried and autoradiographed for 24–28 h at –70 °C with an intensifying screen. Lane 1, ethidium bromide-stained gel of pFA76 DNA digested with *Pst*I and *Ava*I. Lane 2, parallel digest hybridized with ³²P-labelled T₁-4 (2.5 $\times$ 10⁵ c.p.m.). Lane 3, parallel digest hybridized with ³²P-labelled T₁-3 (4 $\times$ 10⁵ c.p.m.). Lane 4, ethidium bromide-stained gel of pFO251 DNA digested with *Pst*I and *Eco*RI. Lane 5, parallel digest hybridized with ³²P-labelled T₁-20 (5 $\times$ 10⁵ c.p.m.). The sizes of the restriction fragments in base pairs are indicated.

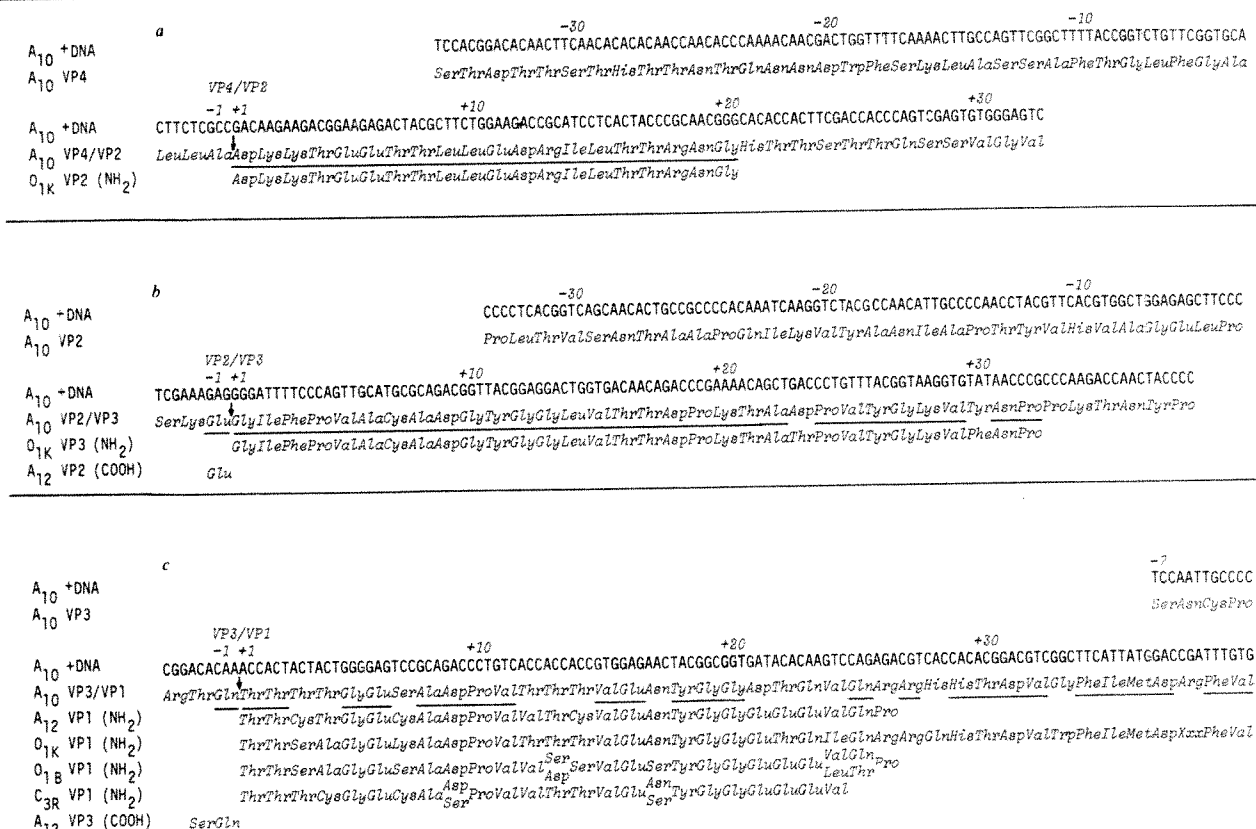


Fig. 3 Nucleotide and amino acid sequences at the junctions of the regions coding for the major capsid proteins. The nucleotide sequence of the positive strand of pFA76 was determined³¹ and used to deduce amino acid sequences in three regions (*a*, *b* and *c*; Fig. 1). The boundaries of the putative coding regions (↓) were set as described in the text. The amino acids are numbered from the amino-terminus of VP2, VP3 and VP1 in *a*, *b* and *c*, respectively. The published carboxy-terminal sequences (COOH) are from ref. 22; the published amino-terminal sequences (NH₂) from ref. 23 (type O_{1K}) and ref. 24 (types A₁₂, O_{1B} and C_{3R}). The underlined residues indicate homology between the A₁₀ sequence and the published data (ambiguities in the published amino-terminal sequences of VP1 are assumed to conform, where possible, to the consensus sequence).

The limited comparisons presented in Fig. 3 are consistent with this hypothesis: the sequence (Thr-Gln) which immediately precedes the amino terminus of VP1, is very similar to the published carboxy terminus of VP3 (Ser-Gln) from a different strain; the difference (Thr rather than Ser) is conservative with respect to charge and hydrophobicity. As the comparison for the carboxy terminus of VP2 is limited to only one residue (albeit identical with the published datum) and there are no comparative data for VP4, a definite conclusion on the contiguity of the structural polypeptides in p88 must await more extensive peptide analysis. Nevertheless, it is clear that the amino acid sequences at three proteolytic cleavage sites within p88 differ in both the type and identity of the amino acids.

VP1 purified from chemically disrupted virus particles has been shown to elicit low levels of neutralizing antibody in pigs and guinea pigs and to protect pigs against the disease²⁵⁻²⁷. An alternative approach to producing a vaccine therefore might be to prepare VP1 from genetically manipulated microorganisms. The recombinant pFA76 contains the sequence encoding all of p88 plus part of both p20 and p52. Such a plasmid offers the option of genetically engineering bacteria so that they synthesize all the capsid proteins in one precursor. This may have advantages over synthesis of VP1 by itself because in the normal replication cycle of FMDV, all capsid proteins are synthesized via the p88 precursor and this may influence their final conformations and immunological properties. Recombinant clones containing the coding sequences of all the capsid proteins, such as pFA76 described here, provide a starting point for such developments.

While this manuscript was in preparation we learned of similar work being done by Küpper *et al.*²⁹ on FMDV strain O_{1K}.

We thank Ruth Ashforth, Sharon Coleman and Cora Paynter for technical assistance and Drs J. Beard, R. B. Flavell, A. Newman, D. Sherratt and R. Thompson for gifts of materials and/or advice.

Received 1 December 1980; accepted 9 March 1981.

1. Chatterjee, N. K., Bachrach, H. L. & Polatnick, J. J. *Viol.* **69**, 369-377 (1976).
2. Sangar, D. V., Rowlands, D. J., Harris, T. J. R. & Brown, F. *Nature* **268**, 648-650 (1977).
3. Rowlands, D. J., Harris, T. J. R. & Brown, F. *J. Virol.* **26**, 335-343 (1978).
4. Sangar, D. V., Black, D. N., Rowlands, D. J., Harris, T. J. R. & Brown, F. *J. Virol.* **33**, 59-68 (1980).
5. Sangar, D. V. *J. gen. Virol.* **45**, 1-13 (1979).
6. Wild, T. F. & Brown, F. *J. gen. Virol.* **1**, 247-250 (1967).
7. Wild, T. F., Burroughs, J. N. & Brown, F. *J. gen. Virol.* **4**, 313-320 (1969).
8. Emtage, J. S., Catlin, G. H. & Carey, N. H. *Nucleic Acids Res.* **6**, 1221-1240 (1979).
9. Hoeijmakers, J. H. J., Borst, P., Van den Burg, J., Weissmann, C. & Cross, G. A. M. *Gene* **8**, 391-417 (1980).
10. Twigg, A. J. & Sherratt, D. *Nature* **283**, 216-218 (1980).
11. Lederberg, J. & Cohen, S. *N. J. Bact.* **119**, 1072-1074 (1974).
12. Boyer, H. W. & Roulland-Dussoix, D. *J. molec. Biol.* **41**, 459-472 (1969).
13. Eckhardt, T. *Plasmid* **1**, 584-588 (1978).
14. Harris, T. J. R., Robson, K. J. H. & Brown, F. *J. gen. Virol.* **50**, 403-418 (1980).
15. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
16. Hanahan, D. & Meselson, M. *Gene* **10**, 63-67 (1980).
17. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965 (1975).
18. Birnboim, H. C. & Doty, J. *Nucleic Acids Res.* **7**, 1513-1523 (1980).
19. Porter, A. G., Fellner, P., Black, D. N., Rowlands, D. J., Harris, T. J. R. & Brown, F. *Nature* **276**, 298-301 (1978).
20. Harris, T. J. R. *J. Virol.* **36**, 659-664 (1980).
21. Robson, K. J. H., Harris, T. J. R. & Brown, F. *J. gen. Virol.* **37**, 271-276 (1977).
22. Bachrach, H. L., Swaney, J. B. & Van de Woude, G. F. *Virology* **52**, 520-528 (1973).
23. Strohmaier, K., Wittman-Liebold, B. & Geissler, A. W. *Biochem. biophys. Res. Commun.* **85**, 1640-1645 (1978).
24. Bachrach, H. L., Morgan, D. O. & Moore, D. M. *Intervirology* **12**, 65-72 (1979).
25. Laporte, J., Grosclaude, J., Wantyghem, J., Serge, B. & Rouze, P. *C. r. hebdo. Acad. Sci. Séanc. Paris* **276**, 3399-3401 (1973).
26. Meloen, R. H., Rowlands, D. J. & Brown, F. *J. gen. Virol.* **45**, 761-763 (1979).
27. Bachrach, H. L., Moore, D. M., McKercher, P. D. & Polatnick, J. J. *J. Immun.* **115**, 1636-1641 (1975).
28. Doel, T. R., Sangar, D. V., Rowlands, D. J. & Brown, F. *J. gen. Virol.* **41**, 395-404 (1978).
29. Küpper, H. *et al. Nature* **289**, 555-559 (1981).
30. Harris, T. J. R. *Nucleic Acids Res.* **7**, 1765-1785 (1979).
31. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).

BOOK REVIEWS

Newton unravelled

Marie Boas Hall

AS Voltaire remarked 250 years ago, few people read Newton because to do so intelligently one must be learned, but everybody talks about him. All scientists know a little of Newton but few know much, nor has it been easy to unravel the details of his life and work. For well over a century and a half scholars have known that Newton left a great mass of manuscript material; only over the past 30 years has it been studied seriously, transcribed and commented upon; and until now the standard biography has been that of David Brewster, published in 1855. Brewster had access to most of the manuscripts but though the Royal Society tried to make them accessible they were not finally deposited, in part, in the Cambridge University Library until 1888. No subsequent biographer utilized them, although Louis Trenchard More used and partly published some papers retained in private possession in his biography of 1934.

When Professor Westfall began work 20 years ago on this monumental scientific biography, destined to become the standard *Life* for the foreseeable future, he knew he must plunge into manuscript sources, for only Newton's *Correspondence* was in course of publication. As he worked, so the mathematical papers have been published *in extenso* by D.T. Whiteside, while many of the papers relating to the *Principia* and *Opticks*, early notebooks, thoughts on the nature of matter, selections from his historical, theological and alchemical manuscripts and numerous minor works have appeared with scholarly commentary, creating what Westfall, who has himself contributed to it, has called "the Newton industry".

Never at Rest is an able and thorough study of Newton the man and the scientist, especially and inevitably the latter who is both the more enduring and the better documented — almost too well documented. Newton was a compulsive scribbler: he read pen in hand, taking notes, making comments, writing and re-writing, analysing, commenting, organizing, drafting and redrafting, and seldom or never destroying.

Professor Westfall has worked through a great deal of the available manuscript material and presents a detailed account — careful, clear and judicious. First comes an analysis of the intellectual background of the day, then Newton's initiation into it,

Never at Rest: A Biography of Isaac Newton. By Richard S. Westfall. Pp.960. ISBN 0-521-23143-4. (Cambridge University Press: 1980.) £25, \$49.50.



Isaac Newton, man of many talents

Mary Evans

then his first steps towards original thought (in his reaction to Descartes' philosophical position) and his very first attempts at mathematics. Soon, leaping over conventional Euclidean geometry, Newton was deep into and going beyond contemporary developments in what we call analysis and he called "resolving problems by motion" (1665) and "the method of tangents", and "analysis by infinite series". In 1669, he showed a paper on the last subject (*De analysi*) to Isaac Barrow, first Lucasian Professor of Mathematics. Barrow immediately recognized its worth and when John Collins (a mathematical news monger, remembered for his

collection of contemporary letters published by Rigaud in 1841) sent him a new book which touched on series, Barrow in return sent *De analysi* (whence, in fact, news of the rising mathematician reached the Royal Society and through its secretary, Oldenburg, passed to Continental mathematicians). Had either Barrow or Collins been more adroit, *De analysi* would have been published and much later controversy avoided. When Barrow soon after resigned his professorship, Newton succeeded him; Westfall agrees with the tradition that Barrow's influence was here paramount.

Although Newton continued to work on and off at mathematics, he lectured on optics in which he had been interested for some years. The Latin lectures were deposited in the University Library as the regulations required, to be published only posthumously, but they became also Book I of the English *Opticks* in 1704, dealing with prismatic phenomena and dispersion. Soon Newton was led to the invention of the reflecting telescope, and of this, as he did not with mathematics, he wrote and spoke freely. The Royal Society learned of it in December 1671, elected him a Fellow, and with his permission both published an account and sent a written description to the Continent. Soon Oldenburg tactfully extracted Newton's first optical paper, with permission to publish it in the *Philosophical Transactions*. Although the Society praised the paper, it was by no means universally approved: Hooke, the Society's Curator of Experiments, an established scientist whose views Newton had earlier studied in *Micrographia* (1665), was acrimoniously and rudely critical; from Paris, Pardies (later a convert) and Huygens were politely so. Newton patiently answered their objections and in 1675, in the midst of controversy with the English Jesuits at Liège, sent his "Hypothesis on light and colours", whose publication he did not then permit. Yet he never blamed Oldenburg; indeed, after the latter's death in 1677, Newton told Hooke that he intended to publish his optical papers as Oldenburg had desired. Some sheets were printed, but the book never appeared — Westfall believes because of accidental destruction by fire.

Meanwhile Newton's interest was elsewhere — in chemistry or alchemy, an interest first aroused by his reading of

Boyle, and then increasingly by the avid study of the vast literature of alchemy, on which he made incredibly extensive notes. He was an able experimenter, so much his notebooks reveal. There is no doubt that he thought alchemists rational, and believed that their strange language could be mastered. Evidently he believed that he understood alchemical symbolism; unfortunately he never translated it, and the key to his experimental programme is thus lost. Westfall believes that Newton's theory of attractive force was the outcome of this work — a debatable view. Newton never completed any papers on chemistry, and only one fragment was published in his lifetime, an incomplete theory of acids in the *Lexicon technicum* (1710) of John Harris. (Westfall unaccountably overlooked this publication, though he knew of other contemporary copies.) Many of Newton's views appear in drafts for a conclusion to the *Principia*, and were added to the *Queries* in *Opticks* in 1706 and 1710, in abbreviated form. At this time Newton also turned to theology, seeking for "primitive" Christianity and becoming a convinced, though secret, Arian.

But during these years he continued to work occasionally on mathematical and mechanical problems. Thus it is not very surprising that he could readily rise to Halley's challenge in 1684, to take up the problem of the Solar System in deep mathematical detail and produce, first a treatise *De motu* and then, under the Royal Society's encouraging praise and Halley's tactful prodding, the triumph of the *Principia* (1687). Westfall discusses the development of the work at some length, though there is more detail in the later chapter on Newton's revisions for the second edition (1713), when Newton was still working on major problems in old age, though less brilliantly.

With the publication of the *Principia*, whose importance was universally recognized, Newton's productive scientific life came near to an end. Perhaps, at 45, he knew it; at any rate, like many scientists he turned to other things. He was active in university affairs in 1688 and Member of Parliament for the University. He began to look for outside preferment, which he did not at once find. He suffered a brief mental breakdown in 1693. Westfall, totally rejecting Manuel's Freudian interpretation, for sound reasons, and equally the recent suggestion of mercury poisoning, ascribes it to a complex variety of factors, not least the mental exhaustion produced by the creative achievement of the *Principia*. Newton recovered, secured his job in London in 1696 as Warden (and later Master) of the Mint, and filled the posts with energy and distinction. In 1703 he became President of the Royal Society, here also showing both administrative ability and energetic dominance. Inevitably, he embarked on no new scientific work, although even in his seventies he was still capable of skilful application when

stimulated. He put *Opticks* into print, and revised it. He worked with Cotes on the second edition of the *Principia* from 1709 to 1713, while still busy at the Mint, and actively engaged in the vituperative dispute with Leibniz over the invention of the calculus. (Each refused to believe that the other had independently developed it, and each was driven on by hot-headed disciples.) The consequences of Newton's earlier failure to publish came home to roost. In extreme old age he turned again to theology, and to Biblical chronology; as with alchemy he sought reason and logic in mystical writings. He became the leading scientist of the first quarter of the eighteenth century, dominating optics, mechanics and mathematical astronomy; only the new mathematics owed nothing to him.

What of Newton the man? The biographical details are meticulously set out here as far as they exist: his family origins, his relations, his schooling, his life at Cambridge and in London. I find Westfall less successful in reconstructing Newton the man than Newton the scientist, whom he clearly prefers. Many of Newton's less agreeable characteristics were common in his day: his reluctance to publish and jealousy over priority were

shared by Huygens, as well as by many others. Was he really so isolated at Cambridge as tradition suggests? He got preferment early, and all his life he attracted and kept the loyalty of many young men. In the priority dispute he seems pathologically paranoiac; so does Leibniz; neither was attractive. In his old age Newton was clearly domineering, easily provoked to anger, impatient of lesser minds, vain and unjust to adversaries — he was never Voltaire's passionless hero. But neither was he totally and disagreeably ungenerous in his dealing with others, especially in his prime.

Never at Rest will be warmly welcomed by Newton scholars, although they will inevitably dispute some points. (Some may feel that Westfall in sticking so close to manuscript sources has been less than generous to earlier labourers.) The book is not easy going, but it provides a masterly, well-documented summary of contemporary views of all the many facets of Newton's astoundingly wide-ranging career which will be novel to all but the specialist, and will be essential reading for aspiring Newtonian scholars. □

Marie Boas Hall retired as Reader of History of Science and Technology at Imperial College, University of London, in 1980.

America and China: lessons from the past

Douglas P. Murray

American Science and Modern China 1876-1936. By Peter Buck. Pp.283. ISBN 0-521-22744-5. (Cambridge University Press: 1980.) £15, \$24.95.

THE recent and rapid embrace of Americans and Chinese has proceeded with little reflection on the troubled history of their cultural interactions. Inherent in this heavy but rewarding book is a call for caution lest we be fated "to repeat elsewhere the failure of American science in modern China". Not only elsewhere, but perhaps again in China itself.

Although Peter Buck describes his book as "an essay in comparative history", he is concerned far less with comparison than with the interaction of two societies that seemed destined to miss each other. He analyses the motives, problems and impact of those pioneers of modern science involved in the cross-cultural transmission of ideas and institutions during the late nineteenth and early twentieth century. Their failure is Buck's fundamental thesis, judged upon the stated goals of America's scientific missionaries and the view that modern science through the 1930s remained "wholly marginal to Chinese life". Yet in the sweep of his complex social history the criteria for what might have been success remain obscure; the legacy of American science that clearly is influencing

the People's Republic of China today is generally ignored.

As reflected in 25 pages of footnotes, the volume is meticulously researched, drawing liberally from, but not duplicating, the important work of other recent analysts of the West's educational impact in China. Buck's focus is somewhat different and original: the "transfer of scientific ideas and institutions to China primarily for what it shows about the United States, and only secondarily because of its bearing on Chinese scientific development" (p.1).

In a brief introduction, Buck assesses our own nineteenth-century debates about the proper role and social power of science and medicine, and his first chapter traces how these conceptions were developed and then acted out in China through medical missionaries. Chapters 3 and 4 focus, respectively, on the subsequent evolution of "laboratory science" (seen by one writer of the time as a "retreat . . . far removed from the strife of sects or wayward passions of the day") and its expression in the China Medical Board and Boxer Indemnity Scholarship program, and on the social origins and orientations of the US-trained scholars who in 1914 founded the Science Society of China.

Buck then returns to the United States to investigate the role of Cornell University

which by 1910 was, with Harvard, "enrolling more Chinese than any other American college" and, indeed, was the birthplace of the Science Society of China. As in the preceding chapter, the author-historian turns sociologist and here covers the most interesting ground. Viewing the development of American science in the context of a society undergoing industrialization, he demonstrates how Chinese were trained in the US for a world that would not exist for them at home and how "in formulating their plans for scientific development, the Science Society's leaders filtered out much of the social content with which American scientists had invested their images of research and laboratory science" (p.164). But, for Buck's purposes, the more critical point comes later, in Chapter 6: the result only further underscored how thoroughly interpretations of these ideals varied, even in the United States . . . In industrial America, Chinese scientists saw integration and order, whereas in reality there was intense conflict and turmoil . . . their image of American industrialization owed more to the ideals of American scientists than to the actual American experience (p.206).

American conceptions of science tended to reach China after they already had changed or been challenged at home.

In a concluding chapter covering the 1920s and 1930s, Buck reviews the belated awakening of China's scientists to their dilemmas both in making science socially relevant and in coming to terms with the growing disappointments of their American benefactors. Yet, ironically, Chinese scholars also were becoming "progressively more convinced that the proper audience for their scientific work lay outside their country", with the international community (p.224). Making a rare link to more recent times, the author notes Mao Zedong's complaint about the behaviour of China's scientists before 1949:

He understood that they were not hostile or indifferent to the Chinese revolution, but from his perspective their enthusiasm for science involved a distinctive evasion of reality; [they mistakenly] 'stayed away from politics' (p.231).

Buck suggests a similar message for the West: detached from social context and public purpose, the best of science can be irrelevant. Our experience with China is the tapestry on which his message is woven.

In principle, this book should be required reading for all in the West now involved in or contemplating academic or technical cooperation with the PRC. Unfortunately, in practice it probably will appeal to a limited audience of scholars — those partial to the complex argumentation of an historian-philosopher whose commitment to grand concepts, intellectual rigour and direct quotation from a myriad of sources often makes the story hard to follow, or at least to enjoy. □

Douglas P. Murray is Director of the US China Relations Program at Stanford University.

Catching up with synchrotron radiation

Yves Farge

Synchrotron Radiation Research. Edited by Herman Winick and S. Doniach. Pp.754. ISBN 0-306-403-63-3. (Plenum: 1980.) £40.95, \$65.

THE past decade has seen an exponential growth in the use of synchrotron radiation — emitted by relativistic electrons in the bending magnets of storage rings — as an experimental probe. A consequence of this rapid advance is that, as in so many other fields, a book on the subject is already dated by the time it appears; this volume is no exception. Many of the 22 contributions will appear obsolete to the specialist, yet they do give a clear view of progress in synchrotron radiation research and of how its application is spreading from physics into chemistry, biology and applied sciences.

The first and last contributions are devoted to a discussion of how photons are generated by electrons passing through bending magnets. This radiation can range in wavelength from infrared to X-rays, is of very high intensity and its pulsed form can be used. Other advanced ideas, such as the use of wiggler or undulator magnets, which are now being tested on storage rings, and the development of a free-electron laser using undulators on these machines, are also discussed. In the remainder of the first part of the book, vacuum, ultraviolet and soft X-ray experiments on atoms, molecules and solids are considered; topics covered include absorption, fluorescence (radiative lifetimes and time-resolved spectroscopy) and photoelectron spectroscopy. These chapters give the reader a good view of the different kinds of electronic excitations which can be observed in such systems. One flaw, however, is that the chapter on photoelectron spectroscopy in solids is incomplete in that angle-resolved analysis of photoelectron spectra is largely ignored; even in 1978, from when most of the contributions date, this technique was known to be an important tool for band-structure determination in solids.

Of the later contributions, those on X-ray absorption spectroscopy are particularly good: Kronig's oscillations or EXAFS can provide useful structural information about a specific chemical element in inorganic as well as organic or biological samples. For example, study of the detailed structures of absorption edges reveals information about ionicity and coordination of the absorbing element.

The chapters on X-ray scattering provide an introduction to a field which was in its infancy at the time of writing (small-angle X-ray scattering, protein crystallography); they are mainly concerned with biological applications. Discussions of various applications, such as topography, X-ray fluorescence analysis, X-ray imaging and microscopy, complete the book.

This volume will be useful to non-specialists, as well as those already using synchrotron radiation in their research. One drawback is that work carried out at Stanford is emphasized at the expense of that performed in Europe, the USSR and Japan. However, the book clearly illustrates the revolutionary implications of the introduction of the technique into the natural sciences and is a cogent argument for the need for new machines designed specifically for the production of synchrotron radiation, some of which are under construction in several countries. □

Yves Farge was formerly Director of the French Synchrotron Radiation Facility (LURE), and is Chairman of the Synchrotron Radiation ad hoc Committee of the European Science Foundation and Scientific Counsellor at the Secrétariat d'Etat à la Recherche.

Reactions on a leash

Johannes Meienhofer

Polymer-Supported Reactions in Organic Synthesis. Edited by P. Hodge and D.C. Sherrington. Pp.484. ISBN 0471-277-12-6. (Wiley: 1980.) £29.50, \$88.

SOLID-PHASE synthesis of peptides and, recently, of oligonucleotides has received widespread attention, due in part to the interesting biological activities of these substances. Somewhat less notice has been given to the development of an impressive variety of polymer-supported organic chemical reactions in which a reagent, a catalyst or a substrate is attached to a polymer. There was a need to gather all aspects of polymer-supported organic synthesis into one volume, a task that the editors have accomplished admirably. A multitude of interesting reactions in many areas of chemistry is presented in 10 chapters, each written by an expert.

The opening contribution by D.C. Sherrington provides the reader (including the novice) with a solid base for understanding both the nature and the potentials of solid-support systems, in particular of the polystyrene type. P. Hodge then considers polymer-supported reagents. Oxidations, reductions and substitution reactions, condensing and acylating systems, solid phase ylids and organometallic reagents are presented. The large amounts of data are clearly organized in 42 well-designed tables and discussions of reaction mechanisms are aided by many schemes, equations and structures.

Sherrington discusses catalysis by ion-exchange resins and related materials, including a survey of transfer catalysis. Opportunities and constraints of catalysis

by enzyme-like polymers are covered by T. Kunitake, and those of transition metals by C.V. Pittman Jr. Chapter 6, on synthesis using polymer-supported protecting groups, by J.M.J. Fréchet, includes asymmetric synthesis and preparation of alkaloids, unsymmetrical carotenoids and other interesting syntheses. A sober and thorough assessment of the present scope and limitations of polymer-supported synthesis in general concludes this chapter. The contributions of J.M. Stewart, Fréchet and M.J. Gait describe solid-phase synthesis of peptides, oligosaccharides and oligonucleotides, respectively. Although the peptide and nucleotide topics have been reviewed frequently and in more detail elsewhere, these three skilfully prepared chapters are well worth reading. Solid-

phase oligosaccharide synthesis still needs much work to become competitive. Completing the main text is another chapter by Hodge, on various applications of polymer-supported species. Rounding off the book is an appendix which provides experimental guidance on the art of suspension polymerization, and on chemical modification of polymers.

This useful book gives thorough attention to detail and provides a wealth of clearly drawn illustrations. Research into polymer-supported reactions is a young and growing enterprise. This book should certainly further its development. □

Johannes Meienhofer is in the Chemical Research Department of Hoffmann-La Roche, Inc., Nutley, New Jersey, USA.

Of plants and pathogens

John Friend

Plant Disease: An Advanced Treatise. Vol. 5, How Plants Defend Themselves. Edited by James G. Horsfall and Ellis B. Cowling. Pp.534. ISBN 12-356-4050. (Academic: 1980.) £27.80, \$49.50.

PLANTS have a range of natural defences which enable them to grow successfully and produce healthy crops despite the widespread presence of many potential pathogens. These resistance mechanisms have had an impact on plant evolution and further understanding of the processes is required in order that the mechanisms can be utilized and adapted by plant breeders and agronomists to improve crop production and yield.

In *How Plants Defend Themselves*, the final volume of a multi-author treatise on plant disease, the various ways in which plants resist disease are covered in 23 chapters. The early sections cover well-documented resistance phenomena such as escape from, and tolerance to, disease.

Later contributions deal with preformed barriers to attack, both physical and chemical. Physical barriers may be external, such as specialized types of plant surfaces and their coverings, or internal, including suberized and lignified tissues, gums, resins and silica. Preformed defensive chemicals include relatively simple compounds exuded from roots and specific antifungal and antibacterial compounds, often of chemotaxonomic interest, present inside the plants.

Considerable effort is now being expended on research into active mechanisms of plant resistance to pathogens, a topic which is covered in considerable depth in the book. Recognition and compatibility between host and parasite which are the most important initial events are surveyed in detail by Sequeira, who pays particular

attention to the possible functions of lectins, polysaccharides and lipopolysaccharides. Király reviews hypersensitivity with particular reference to his own recent experiments, and concludes that it is not necessarily the cause of resistance but probably a result of stress that develops in incompatible host-parasite combinations. Following a review of active formation of physical barriers by Beckman, Cruickshank provides a thoughtful review of the phytoalexin concept and the possible roles of phytoalexins in disease resistance.

However, the highlight of the book is a series of five chapters on acquired immunity to a second attempted invasion triggered by an initial infection by viruses, bacteria, fungi, nematodes and insects. In the main these chapters are descriptive, but the two on viruses and bacteria discuss possible mechanisms for this type of resistance which could involve alterations in nucleic acid and protein metabolism.

The phenomena are of great importance for an understanding of resistance mechanisms and may also have practical application in the design of new fungicides.

The book is an excellent summary of the most recent work on plant disease resistance. It can be criticized because it under-emphasizes the stimulus given by plant pathologists to research on the biochemistry and physiology of healthy plants; nevertheless, the reader should realize that the editors deliberately tried to produce a series in which the authors would write reflective articles examining the future of research on plant disease. Despite the rather anthropomorphic chapter headings and the idiosyncratic suggestion that plant resistance is analogous to the defence of a mediaeval castle, this editorial policy has been successful. □

John Friend is Head of the Department of Plant Biology at the University of Hull.

Nature's archives

Robert Olby

Natural History Manuscript Resources in the British Isles. Compiled by G.D.R. Bridson *et al.* Pp.473. ISBN 0-7201-1559-0. (Mansell, London/Bowker, New York: 1980.) £97.

EVER since the establishment and consolidation of local government authorities in the nineteenth century, archive material has been preserved in provincial record offices, public libraries and civic universities on an increasing but uncoordinated basis. Successive parliamentary acts have expanded the list of official repositories for public documents, most of which are held either in a record office under the jurisdiction of the chief executive (formerly the town clerk) or in a public library under the control of the librarian. The scattering of archives between public libraries, record offices, museums, local societies, professional institutions and universities is inevitable given the varied sources generating them and the diversity of the parties interested in studying them. That the reform of local government in 1974 failed to clarify the relationship between the record office and public library is hardly surprising. That there should be some rivalry between civic universities and public libraries over the deposit of documents is to be expected. The important problem is that of guiding the inquirer in his search for source material.

There have been a number of initiatives to offer such guidance and to stimulate the formation of collaborative policy for deposition. The Commission on Historical Manuscripts, which celebrated its centenary in 1969, has played an invaluable role through the National Register of Archives which was established in 1949. Of particular importance for the history of science is the work of the Contemporary Scientific Archives Centre at Oxford.

The initiative for the present work goes back to the 1960s when D.E. Allen, R.H. Ellis and others formulated an ambitious plan to conduct a major search of the country's archives which relate to natural history. In the event adequate funds were not forthcoming and a more modest survey was undertaken by Bridson, Phillips and Harvey. They are to be congratulated for the extent and the detail of their book. There have always been close links between botany, zoology and natural history. Since the beginning of this century experimental biology has distanced itself from the work of the naturalists, but for most of the period covered by this survey students should find this book an aid to research. The only regret is that the authors were unable to include the extensive collections of the Public Record Office at Kew. □

Robert Olby is Reader in the History of Science, Department of Philosophy, University of Leeds.